ECOLOGICAL, PHENOTYPIC AND GENETIC DIVERSIFICATION IN *Oophaga* POISON FROGS

A Thesis Submitted to the

College of Graduate Studies and Research

in Partial Fulfilment of the Requirements

for the Degree of Doctor of Philosophy

in the Department of Biology

University of Saskatchewan

Saskatoon

By

ANDRES POSSO-TERRANOVA

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Biology,
University of Saskatchewan, 112 Science Pl.
Saskatoon, Saskatchewan (S7N5E2)

ABSTRACT

Despite the incredible diversity of lowland tropical rainforests; we still have limited understanding of the drivers of speciation in these ecoregions. In these difficult-to-access habitats that maintain a significant reservoir of the world's biodiversity, it is widely accepted that a vast amount of species would remain undescribed and that some of them will likely go extinct before they have a chance to be properly described and studied. The present work studies the ecological, phenotypic and genetic diversification of Neotropical Harlequin poison frogs of the genus *Oophaga*, a group of organisms where a perplexing variation in colouration and patterns have led to taxonomic uncertainty. First, in order to investigate the relative contribution of geographical and environmental factors to the diversification of these frogs, I combined phylogenetic methods with detailed geographic data and environmental niche modelling (ENM) to test the role of geographic isolation, climatic niche divergence and altitudinal gradients. Overall, my results suggested that speciation along climatic gradients on a structured landscape has been a major evolutionary force behind the diversification of *Oophag*a poison frogs. Second, by using an integrative taxonomy framework and different lines of evidence derived from environmental data and intrinsic biological attributes (phenotypic variation, genetics), I statistically tested competing lineage-boundary hypothesis within the currently known *Oophaga histrionica* species. I found that diversity within the complex has been underestimated and proposed the existence of at least three independent evolutionary lineages of *Oophaga* poison frogs that should be further studied as they potentially represent new species to science. These results have important conservation implications as some of its members are considered amongst the most endangered species of all amphibians. Third, I hypothesized that the genes, pathways, and/ or gene networks potentially associated with colouration, alkaloid metabolism, transport and

storage, should be highly expressed in skin tissue. Then, by taking advantage of the high coverage offered by Next Generation sequencing (NGS), I investigated the common transcriptional profiles of five *Oophaga* lineages. To my knowledge, this represent the first transcriptome dataset for Dendrobatid frogs where a functional annotation and comparative analyses allowed the identification of potential candidate genes in important adaptive traits. The identification of more than 250 orthologous contigs across lineages allowed testing the current phylogenetics hypothesis for this group. Overall, in this part of my work I provide an important molecular resource not only for the study of aposematism within Dendrobatids but for the future assembling and annotation of Dendrobatid genomes. Finally, Harlequin poison frogs represent an excellent group, yet underexplored, to study the ultracellular structural basis of colour variation and the role of candidate genes in colour variation. In this final part of my dissertation, I report the basis of histological-cellular components of colouration and provide initial hypothesis of the functional effect of mutations at the Melanocortin-1 receptor gene (MC1R) and the genetic mechanisms of colour variation in this frogs. One of the most interesting findings was the observation of similar phenotypes appearing independently in nature as a consequence of different mutations of the same gene, a classical example of convergent phenotypic evolution and genetic repeatability. Broadly, I provided evidence to suggest that variation at the MC1R receptor could potentially be a major factor responsible for the high phenotypic variation of aposematic signals of Harlequin poison frogs.

ACKNOWLEDGMENTS

First, I would like to thank Dr. Carlos Carvalho, Dr. Chris Todd and Dr. Yves Plante for their advice as members of the PhD committee. I immensely would like to thank Professor Dr. José Andrés for his invaluable help and support during all of these years. None of this work would had been possible without his mentorship. I couldn't have had a better PhD supervisor.

Many people contributed with technical and administrative support. Many thanks to Halyna Heisler, Deidre Wasyliw, Joan Virgl, Gillian Murza, Dr. Professor Ken Wilson, and specially, Pamela Hind.

I also would like to thank the members of Rick Harrison's lab at Cornell University and Dr. Stefan Lötters for their comments on early versions of our manuscripts. Finally, I would like to highlight that this research was mainly supported through the NSERC-discovery grant # 409354 "Genomics of Speciation" to Professor Dr. José Andrés and a graduate scholarship to Andrés Posso-Terranova through the Department of Biology at the University of Saskatchewan.

Dedication

Parental care, as early stated by Ronald Fisher (*"The Genetical theory of Natural Selection", 1930*), involves any parental investment that benefits the offspring but reducing the parents' ability to invest in other components of fitness. Being far from you was my investment to provide the resources that you still need. I had to do it, and has been painful. I am not sure if I will be able to recover so many things I've missed in your lives…

…For you to remember, you guys are, and will be, the reason of my existence.

To my Children, Andres Mauricio and Maria Jose.

To my parents, Noralba and Manuel. I am what you taught me to be. *To my only brother, Jose Manuel.* We are different, but we are the same. You have been there in all of those moments when I needed you.

Finally, *To the love of my life, Martha L*. Probably you don't know this, but sometimes, you are my legs. Sometimes, you are my arms…

…You are my other half.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AL	Average Transcript Length
AMNH	American Museum Of Natural History
ANOVA	Analysis Of Variance
ARC	Age-Range Correlation
BDA	Blast Descriptor Annotator
BFM	Bright Field Microscopy
BI	Bayesian Inference
BIC	Bayesian Information Criterion
BP	Biological Process
cAMP	Cyclic Adenosine Monophosphate
CAX	Cation Exchanger Proteins
CC	Cell Component
CEG	Core Eukaryotic Genes
CEGMA	Core Eukaryotic Genes Mapping Approach
cv	Carotenoid Vesicle
CYP	Cytochrome P450 Enzyme
DAPC	Discriminant Analysis Of Principal Components
DSS	Difference Of Sums Of Squares
DTT	Disparity Through Time
EL	External Loop
ENM	Environmental Niche Modelling
ER	Effective Number Of Reads
ESS	Effective Samples Sizes
ESU	Evolutionary Significant Unit
GIS	Geographic Information Systems
GLC	General Lineage Concept
GO	Gen Ontology
GPCR	G-Coupled Proteins
GPS	Global Positioning System
GYMC	Generalized Mixed- Yule-Coalescent
HMMs	Hidden Markov Models
HTX	Histrionicotoxin
ICZN	International Code Of Zoological Nomenclature
IRSNB	Institut Royal Des Sciences Naturelles De Belgique
LSID	Life Science Identifiers
MC1R	Melanocortin-1 Receptor
MCC	Maximum Clade Credibility
MCH	Melanin-Concentrating Hormone

MDC	Minimize Deep Coalescence
MDC	Minimize Deep Coalescence
MDI	Disparity Index
MF	Molecular Function
ML	Maximum Likelihood
ms	Melanosome
MU	Management Unit
NBF	Neutral Buffered Formalin
NCBI	National Center For Biotechnology Information
NE	North-East
NEF	Nikkon Electronic Format
NGS	Next Generation Sequencing
NIH	National Institutes Of Health
NMDS	Non-Metric Dimensional Scaling
NW	North-West
PCA	Principal Component Analysis
PCA	Principal Component Analysis
PCNM-RDA	Principal Coordinates Of Neighbour Matrices
PCoA	Principal Coordinates Analysis
PERMANOVA	Permutational Analysis Of Variance
PNO	Phylogenetic Niche Occupancy
PP	Posterior Probability
pt	Pterinosome
SCN	Single Copy Nuclear
SEEVA	Spatial Evolutionary And Ecological Vicariance Analysis
TCS	Statistical Parsimony Network
TEM	Transmission Electron Microscopy
TIFF	Tagged Image File Format
TM	Transmembrane Domains
TPM	Transcripts Per Million
TPM	Transcripts Per Million
UCLN	Uncorrelated Log-Normal
VEI	Equal- Covariance And Diagonal Distribution Model
VGIC	Voltage-Gated Ion Channel Proteins
α-MSH	Melanocyte-Stimulating Hormone

CHAPTER 1

GENERAL INTRODUCTION

1.1 The importance of geographic variation

"... if variations useful to any organic being ever do occur, assuredly individuals thus characterized will have the best chance of being preserved in the struggle for life; and from the strong principle of inheritance, these will tend to produce offspring similarly characterized. This principle of preservation, I have called, for the sake of brevity, Natural Selection ".

Charles Darwin.

Ever since Darwin, the concept of geographic variation has been intimately related to that of biological diversity and the process of speciation (Mayr, 1963; Futuyma & Mayer, 1980). The modern evolutionary synthesis (Huxley, 1958) states that the main mechanism of change is natural selection over a variety of phenotypes in their surrounding environment and that even slight advantages are important when continued. As Stephen J. Gould put it "*the foundation of most evolutionary theory rests upon inferences drawn from geographic variation or upon the verification of predictions made about it*" (Gould & Johnston, 1972), implies that geographic variation retains a central role in our understanding on the nature of species and speciation.

Geographic variation in phenotypic traits is ubiquitous, the literature on this topic is immense and many perplexing examples can be found in all major groups of vertebrates (Peterson *et al.*, 2002; Illoldi-Rangel *et al.*, 2004; Rosenblum *et al.*, 2004; Amézquita *et al.*, 2011; Wang, 2011; Corso *et al.*, 2012). This variation has been explained by the balance of gene flow and selection when a species has a wide distributional range relative to its dispersal capacity (Jang *et al.*,

2011). When populations are geographically isolated, phenotypic differences may lead to intra-specific divergence due the effects of reduced gene flow and the random variation of alleles frequencies by genetic drift (Dobzhansky, 1940; Chiucchi & Gibbs, 2010; Chouteau & Angers, 2012); however, population divergence could also be driven by diversifying selection in sympatric populations (Smith, 1966; Nokelainen *et al.*, 2012; Valkonen *et al.*, 2012). If intra-specific variation is driven by natural selection, a strong association between morphology and the measures of environment that influence adaptation is expected. Such correlations must serve as the primary test for selection in geographic variation (Gould & Johnston, 1972). Overall, the phenotypic variation observed across the distribution of an animal species may depend on genetic drift, natural and sexual selection, or a combination of several of these evolutionary forces (Wu, 2001).

In amphibians, many studies have concluded that gene flow is critical for geographic variation, pointing out that geographic barriers (i.e. barrier effect and isolation by distance) are evident for geographic variation of genetic diversity (Zeisset & Beebee, 2008). Specially in poison-dart frogs (Dendrobatids), colouration and vocalization calls are two well-studied traits due its extreme geographic variation and signalling function for intra-specific communication. Even though a dense mass of literature is available, controversial hypothesis has been proposed for explaining the evolutionary forces leading these traits (Brown *et al.*, 2011). For instance, bright colouration and its variation in Dendrobatids is generally considered to be an example of evolution by natural selection (Noonan & Comeault, 2009; Chouteau & Angers, 2011; Rojas, 2012); however, it can also function in the context of sexual selection (Maan & Cummings, 2008, 2009; Crothers *et al.*, 2011). Along this thesis dissertation, I will focus on Dendrobatid frogs. Broadly, I proposed to study the phenotypic, genetic and ecological diversity across

geography and among different populations/lineages of poison frogs to disentangle these unresolved questions and to shed light in the study of variation in this system (Lougheed *et al.*, 2006).

1.2 Aposematism

When protection and avoidance of potential predators is accomplished by combining a defence mechanism (e.g. toxins or spines) with an easily recognized signal (e.g. acoustic, olfactory or visual), a species is to be called aposematic (Holen & Svennungsen, 2012). The term aposematism was first introduced by Alfred Wallace's friend Edward Bagnall Poulton (Poulton, 1890) who pioneered in providing experimental proofs of the protective value of colouration in insects (Poulton, 1887). The great majority of studied aposematic species use visual warning displays that are often conspicuous in both structure and colour, with internal contrasting colours (i. e. bright red, yellow, orange) in combination with black, blue or brown (Wüster *et al.*, 2004). This association of toxic or distasteful substances associated with a bright colouration has been observed and described in a variety of organism from insects (Svádová *et al.*, 2009; Nokelainen *et al.*, 2012) to molluscs and snakes (Wüster *et al.*, 2004; Becerro *et al.*, 2006) where a potentially costly trait (bright colour) evolved to maximize the effect of another potentially costly trait (toxicity) (Lynn, 2005). It has been proposed that populations of aposematic species should converge into one (or just a few) non-variable and conspicuous phenotypes because variation in aposematic signals makes it difficult for predators to learn and retain the association between colour patterns and distastefulness (Mallet & Joron, 1999; Ruxton *et al.*, 2004; Mappes *et al.*, 2005; Darst & Cummings, 2006; Darst *et al.*, 2006) (Darst *et al.*, 2006). Thus, some authors have considered the spatial variation of aposematic signals, such as that seen in poison frogs (Summers *et al.*, 2003; Grant *et al.*, 2006; Brown *et al.*, 2010; Brown *et*

al., 2011; Crothers *et al.*, 2011; Brusa *et al.*, 2012) as a perplexing paradox. However, mathematical and simulation models suggested that in fragmented environments, random effects combined with anti-apostatic (frequency dependent) selection, low toxicity, low predator and prey dispersal rates result in the establishment of inter-population variation (i.e. polymorphism) in aposematic signals (Sherratt, 2006).

1.3 The uniqueness of polymorphic toxic species

Aposematic animals are usually chromatically monomorphic (Greenwood *et al.*, 1981; Darst *et al.*, 2006). Contrary, colour polymorphism or polychromatism is the occurrence of intra-specific colour variation, at the inter-population but not necessarily at the intra-population level (Hegna *et al.*, 2012; Micheletti *et al.*, 2012). As previously stated, colour polymorphism has been observed in association with toxicity and challenges the phenotypic convergence hypothesis in aposematic signals, suggesting that colour signal variation may serve other purposes or respond to different selective pressures. Among vertebrates, Dendrobatid frogs provide the most significant, if not the only, example of polymorphic species, showing a striking array of colour morphs (Myers & Daly, 1976; Myers, 1982; Lötters, 1992). This perplexing variation is well known and described in the strawberry poison frog *Oophaga pumilio* whose colour diversity spans the visual spectrum including even cryptic variants (Pröhl & Ostrowski, 2011; Wang *et al.*, 2013).

Although the mechanisms by which these polymorphisms are maintained are not yet fully understood, theoretical models suggest that localized positive frequency-dependent selection, if strong enough, could explain the maintenance of geographic mosaics characterized by the aposematic signals being locally monomorphic but displaying differences between localities (Sherratt, 2006). Such selection would lead to situations where predators from different localities

recognize and avoid distinct warning signals; however, empirical studies are required to corroborate this hypothesis (Ratcliffe & Nydam, 2008). Natural selection is not likely to be the only evolutionary force driving colour variation and other factors have been taken into consideration to explain the occurrence of colour polymorphism in aposematic poison frogs species. Other factors include genetic drift (Rudh *et al.*, 2007), sexual selection (Summers *et al.*, 1997; Maan & Cummings, 2009), and variation in toxin availability (Wang, 2011). However, most of this knowledge comes from the studies on the populations of *O. pumilio* found in Bocas del Toro, an small archipelago on the East coast of Panama. Studies in different systems, at different geographic scales are needed before general conclusions can be made.

1.4 Colouration in Anurans

1.4.1 Structural basis

Colour patterns in vertebrates may depend on structural characteristics of integuments or pigments or a combination of both (Endler, 1980). In Anurans, pigments cells are organize within the dermal chromatophore unit, a structure that encompasses three different types of chromatophores (Fig. 1-1). Uppermost in the unit there is a layer of pteridine/carotenoid containing xanthophores. These cells lie above a layer of iridiophores containing crystalline deposits of purines that reflect light, which in turn are underlain by melanin containing melanophores that have finger-like processes that extend up and over the iridiophores (Bagnara *et al.*, 1968). One of the proposed advantages of this unit is that it enables rapid colour changes in response to environmental clues. In tree frogs (*Hyla spp*.) plastic changes in colouration are related to the rearrangements of pigments. When the skin appears light in colour the melanosomes are concentrated in the cell body of the melanophores. When the skin appears dark,

the melanosomes disperse into the melanophore' s processes covering the iridiophores and obscuring the bright colours of the skin (Kraemer *et al.*, 2012). The rearrangement of pigments within cells is not the only in mechanism leading variation in colouration in Anurans. In the fire-belly toad (*Bombina bombina*) the ventral skin does not contain "typical" chromatophore units. In this case, red patches contain only xanthophores and iridiophores while black ones only contain melanophores. These observations suggest that differences in chromatophore composition may also play a key role in colour variation in other species of Anurans. However, no other species have been studied thus far (Rudh & Qvarnström, 2013).

1.4.2 Molecular genetics

While few studies have looked at the molecular genetic basis of colouration in fishes and reptiles (Rosenblum *et al.*, 2004; Gross *et al.*, 2009; Mundy, 2009; Rosenblum *et al.*, 2010), the molecular basis of colouration in amphibians is virtually unexplored and only two published studies have attempted to identify genes related with skin colour variation in this group (Rudh & Qvarnström, 2013). All these studies have focused on the potential role of the melanocortin receptor encoding gene MC1R in the common frog, *Rana temporaria.*

There is ample evidence that the melanocortin receptor encoding gene MC1R is related with colouration in homeotherms (Hoekstra, 2006). In these taxa, MC1R controls which type of melanin is produced in melanocytes. When MC1R is activated, it triggers a series of chemical reactions that stimulate these cells to make blank-brown eumelanin. If MC1R is not activated or is blocked, melanocytes make red-yellow pheomelanin instead of eumelanin. Thus, substitutions reducing the ability of the MC1R to stimulate eumelanin production are associated with light colouration (Domingues *et al.*, 2012).

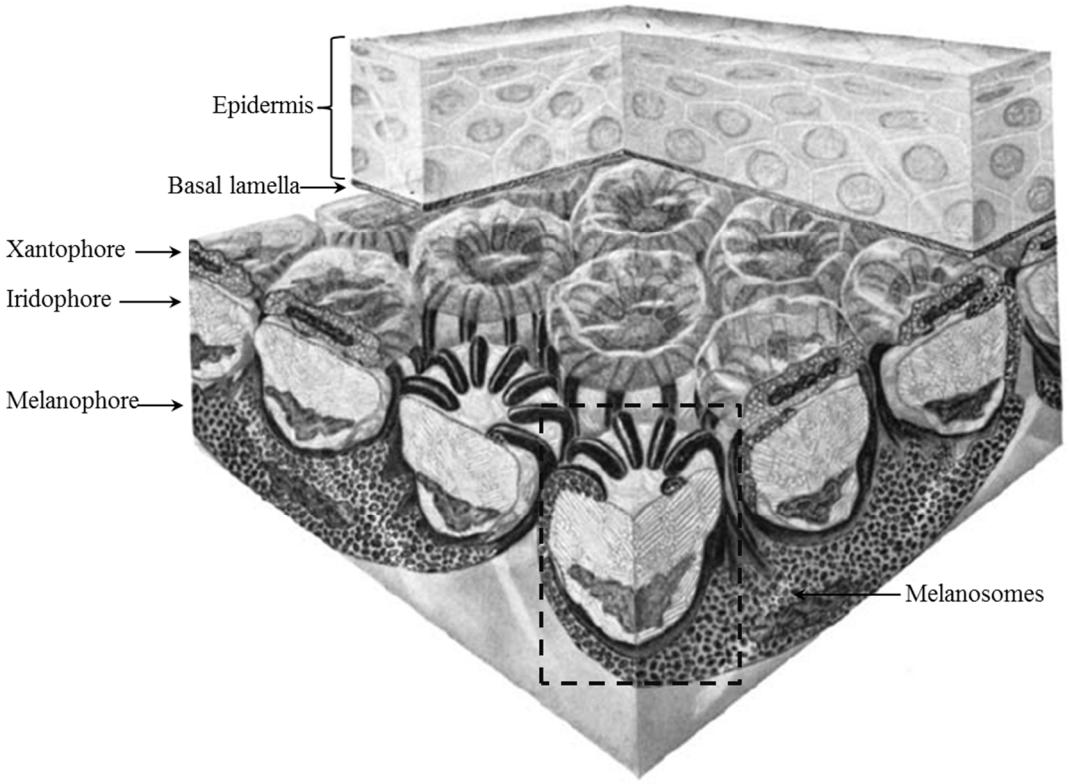


Figure 1-1. Schematic interpretation of the dermal chromatophore unit from several anurans (dotted line). Adaptation to a dark background is represented (Adapted from Bagnara *et al.* 1968).

In comparison with the molecular genetics of melanogenesis in mammals, the regulation of MCR1 in poikilotherm vertebrates is poorly understood (Mangano *et al.*, 1992). However, studies in fishes and reptiles have shown that the activation of the receptor results in changes in the number, size and dispersion patterns of the melanosomes in the melanophores. In the Mexican cave tetra *Astyanax mexicanus*, non-functional alleles are associated with a brown mutant phenotype characterized by a reduced melanin content, a lessened number of melanophores and brownish eyes (Gross *et al.*, 2009). Similarly, in three species of lizards (*Sceloporus undulatus*, *Aspidoscelis inornata* and *Holbrookia maculata*), statistical association between single MC1R mutations and background colouration were highly significant for each

species in individuals adapted to different environments (white sand and dark soil) (Rosenblum *et al.*, 2010).

The only species of amphibian studied so far is the common frog *Rana temporaria*. (Herczeg *et al.*, 2010; Chikako, 2012). Only synonymous changes were found when the MC1R DNA sequence was analysed, neither explaining differences in melanism nor showing evidence of deviation from neutral evolution (Chikako, 2012). However, another colour-related study on *R. temporaria* found a higher degree of melanism with increasing latitude suggesting a thermoregulatory function (Alho *et al.*, 2010). Such a colour change implicates that MC1R could be a key mediator of adaptive colouration in this organisms and is likely to be under selection at particular habitats (Logan *et al.*, 2006).

1.5 The Dendrobatid family: poison-dart frogs

Poison-dart frogs has been used as a common name of a group in the family Dendrobatidae which are native to Central and South America (Myers & Daly, 1976). These species are diurnal and often have brightly coloured bodies. Although all wild poison-dart frogs are at least somewhat toxic, levels of toxicity vary considerably from one species to the next and from one population to another (Darst *et al.*, 2006). These amphibians are called "dart frogs" because the Colombian tribe *Embera* use their toxic secretions to poison the tips of blow-darts for hunting. However, of over 175 species, only three have been documented as being used for this purpose, all of which come from the *Phyllobates* genus, characterized by the relatively large size and high levels of toxicity of its members (Grant *et al.*, 2006). Poison-dart frogs are possibly one of the most charismatic examples of aposematic colouration (Clark *et al.*, 2005; Brown *et al.*, 2010; Richards-Zawacki & Cummings, 2011), and hence, they are an excellent model systems to understand not only the evolution and maintenance of aposematic signals but

also to study the implications of colour pattern variation and its role in the behaviour and ecology. Surprisingly, despite the important role that poison frogs play in evolutionary research and public conservation awareness, very little is known about the demography, population biology and genetics of most species.

1.6 Evolutionary studies of poison-dart frogs in tropical South America

Studies conducted on aposematic frogs of tropical South America showed a complex panorama of the relationship between genetic and phenotypic diversification and the maintenance of phenotypic diversity in nature. For instance, populations of a single Amazonian frog species with low phenotypic variation where found to be divergent when analysed at the molecular level and may be considered a complex of five to nine species (Funk *et al.*, 2011), implying that biodiversity is much more severely underestimated than previously thought. On the contrary, populations of a species may be phenotypically highly diversified despite a lack of concordant genetic divergence as identifiable using nuclear and mitochondrial markers (Amézquita *et al.*, 2011; Brown *et al.*, 2011).

Despite the need to consider relevant phenotypic traits to have a global description of the possible relations between genetic diversity and phenotype, only a few studies have really adopted a comprehensive approach. Recent investigations have focused on reconstructing the phylogeographic patterns (Roberts *et al.*, 2006) and phylogenetic relations within genus as *Ranitomeya* (Brown *et al.*, 2011) and *Oophaga* (former *Dendrobates*) genus (Hagemann & Pröhl, 2007; Hauswaldt *et al.*, 2011). Some studies incorporated data on recent population structure (Solomon *et al.*, 2008; Santos *et al.*, 2009) and almost no scientific studies related to phenotypic traits to biological important components such as advertisement calls and morphology have been conducted in south American tropical anurans (Amézquita *et al.*, 2011).

Within South American species of the genus *Oophaga,* no studies have considered the colouration despite its importance for intra-specific divergence and its role in mate choice at the population level (Brown *et al.*, 2010; Pröhl & Ostrowski, 2011).

Within this framework, it is particularly important to investigate the genetic and phenotypic variation across geographic specie's distribution to have a clear picture of population-level diverging patterns. This will shed light in revealing cryptic diversity where two or more morphologically similar species are erroneously classified under a single species name (Bickford *et al.*, 2007). Together with ecological (i.e environmental) variation, it would be possible to insights into the evolutionary origin of the stunning diversity of phenotypic traits of poison-dart frogs.

1.7 The study area: The Colombian Chocó region

It has been estimated that 60% to 70% of all terrestrial plants and non-fish vertebrates occur just within 1.44% of the Earth's land surface arising the concept of biodiversity hotspot, a biogeographic region with a significant reservoir of biodiversity (Howlett, 2000). High diversity forests occur all over the world but they are concentrated in lowland areas with high and evenly distributed rainfall. Hence, the greatest biological diversity usually occur in northern South American forests (Achard *et al.*, 2002; Carnaval *et al.*, 2009). The biogeographic region known as South American Chocó is considered a diversity hotspot which extension includes eastern Panama throughout the Pacific west lowlands of Colombia to northern Ecuador (Achard *et al.*, 2002).

Specifically, the Colombian Chocó region is one of the most extensive areas of tropical forests in the Neotropics (Mayaux *et al.*, 1998; NOAA, 2013) and is classified as tropical wet and tropical pluvial forest. It holds the highest average annual precipitation record measured at

13,300 mm which makes this region the wettest place in the world with an annual temperature average of 27.3°C (Mayaux *et al.*, 1998). These climatic conditions combined with the thick rain forest and the extreme diversity of living forms make this region an ideal place for conducting ecological and biological studies (Howlett, 2000). However, because of the former political instability of the region, the access to several places is restricted and recent reliable scientific information about fauna and flora are scare or not available.

1.8 The study system: Harlequin poison-dart frogs

The Harlequin poison-dart frogs of the *O. histrionica* complex have an intriguing mating system and extensive parental care. Similar to *O. pumilio*, a sister species, males hold territories which are used as call sites, courtship areas and oviposition sites (Pröhl & Ostrowski, 2011). Males usually call on elevated perches in his territory and the following courtship consists of extended sequences of acoustic, visual, and tactile interactions. Mating lasts on average 1.5 hours and finishes with the female laying eggs and the male deposits sperm usually under a dry leaf (Myers, 1982). The female returns and the tadpoles wiggle up on her back and she transports them to a small body of water (leaf axis). The tadpoles solely feed on unfertilized eggs that the female provide and orphanage tadpoles do not survive. The female deposits unfertilized eggs every day and the tadpole suck out the yolk (Myers, 1982) .

These frogs secrete mainly cardio-toxins known as histrionicotoxins (HTX) through skin glands. These lipophilic alkaloids appear to be sequestered from a diet of alkaloid-containing arthropods (Saporito *et al.*, 2012). The majority of these defensive chemicals present in ants are straight-chain alkaloids, being HTX a major component in myrmicine ants from Central and South America (Saporito *et al.*, 2004; Saporito *et al.*, 2007). These moderate to highly toxic compounds act as potent non-competitive antagonists of nicotinic acetylcholine receptors,

binding to a regulatory site on the delta subunit of the ion channel complex (Tarvin *et al.*, 2016). They also have some affinity for sodium and potassium channels. The artificial synthesis of HTX and various homologues is challenging and has been the subject of many different attempts (Saporito *et al.*, 2012). Besides their toxicity, these frogs intrigue many researchers, not only working in the field of herpetology, but in ecology and evolutionary biology because they show one of the most perplexing colourations of all amphibians and extreme phenotypic variation over their apparently continuous geographic distribution (Myers, 1982). This variation in colours and patterns have led to uncertainty in the taxonomic status and indeed, it has been suggested that different colour morphs within the *Oophaga histrionica* complex may represent several undescribed species endemic to the biogeographic Chocó region (Silverstone, 1975; Myers, 1982; Lötters, 1992; Brown *et al.*, 2011). However, probably because of the remoteness of their location, the patchy distribution of different populations and their increasing rareness, almost no information exists about the phylogenetic relationships between morphs, their taxonomic status or the different factors influencing phenotypic and genetic diversity in this system.

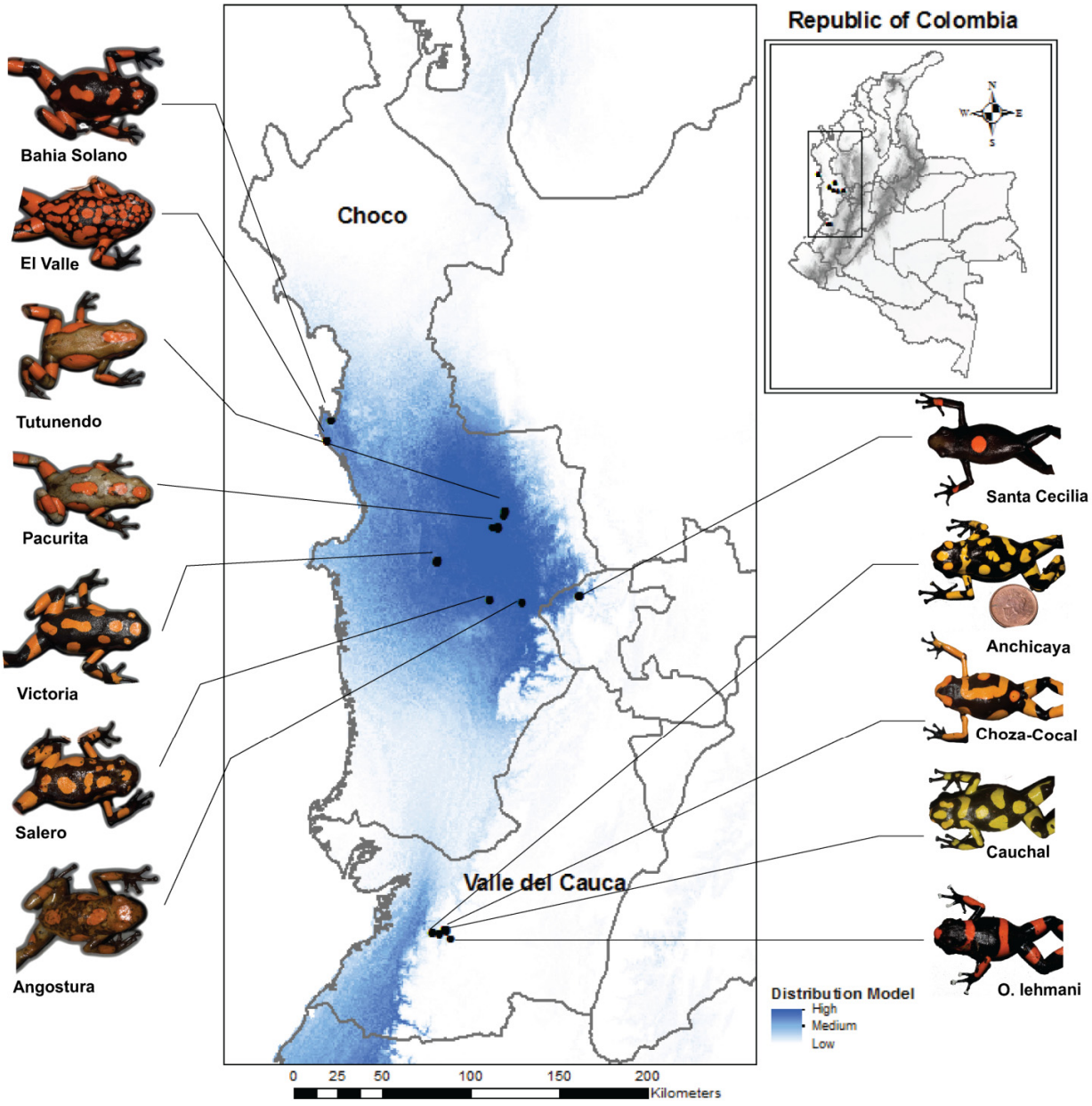


Figure 1-2. Predicted species distribution and phenotypic variation of Harlequin poison frogs. Names on each individual photo indicate the corresponding geographic sampling location (detailed information is presented in the following Chapters).

In this PhD dissertation, I focused on the distribution, geographic variation and population structure of several *Oophaga* lineages, but especially, I focused in members of the *O. histrionica* group (Harlequin poison frogs), a complex of species that is likely to represent several independent evolutionary units new to science. To do so, I used a multi-disciplinary approach that combined morphological, histological, molecular genetics, and geographic information systems (GIS) techniques. Broadly, I applied the mentioned approaches to complete the following objectives:

1. Despite the incredible diversity of lowland tropical rainforests, we still have limited understanding of the drivers of speciation in these ecoregions. I investigated the relative contribution of geographical and environmental factors to the diversification of the Neotropical genus of poison frogs *Oophaga* by combining phylogenetic methods with environmental niche modelling (ENM) approaches.
2. Early taxonomic studies (Berthold, 1846; Myers & Daly, 1976) and preliminary genetic data from geographically scattered individuals (Grant *et al.*, 2006; Brown *et al.*, 2011) have failed to recognized the striking phenotypic diversity and potential ecological divergence of Harlequin poison frogs. I used transcriptome polymorphisms, mtDNA data and microsatellites variation to generate a comprehensive genetic dataset that then, I combined with patterns of phenotypic and environmental divergence. The resulting common signal was used to establish a robust lineage delimitation of Harlequin poison frogs.
3. RNA-sequencing offers an ideal opportunity to further our understanding of the alkaloid sequestration system, auto-resistance to toxic compounds and variation in colouration and pattern in Harlequin poison frogs. By taking advantage of the high coverage offered by Next Generation Sequencing (NGS), (i) I used RNA-sequencing to generate an annotate a *de-novo* skin transcriptome for Harlequin poison frogs; (ii) conducted a comparative transcriptome analysis based on phenotypically variable individuals; (iii) identified potential candidate genes for important adaptive traits and (iv) constructed a genomic-based phylogeny for these polymorphic species with controversial taxonomic status.
4. Harlequin poison frogs represent an excellent group, yet underexplored, to study the

ultracellular structural basis of colour variation and the potential role of candidate genes in colour variation. In the last part of this dissertation, I focused on the histological analysis of skin colour variation and the characterization of the melanocortin-1 receptor gene (MC1R) across the geographic range of several Harlequin poison frog lineages in order to shed light in the structural, genetics and molecular mechanisms of skin colouration.

CHAPTER 2

COMPLEX NICHE DIVERGENCE UNDERLIES LINEAGE DIVERSIFICATION IN *OOPHAGA* POISON FROGS*

2.1 Introduction

The Neotropical region harbours one of the richest amphibian faunas on Earth. Whereas recent large-scale phylogenetic studies have shown that this high tropical diversity in amphibians is likely to be partly related to higher speciation rates in the tropics (Pyron & Wiens, 2013), the evolutionary processes responsible for speciation in the Neotropics continue to engender debate. Multiple biogeographical scenarios have been proposed seeking to understand the origins of diversification including Pleistocene regional refugia, Pliocene-Pleistocene riverine barriers, Miocene-Pleistocene marine incursions, closure of the Isthmus of Panama, etc. (Moritz *et al.*, 2000; Paulay & Meyer, 2002; Hubert & Renno, 2006; Brumfield & Edwards, 2007). Among them, a prevailing set of hypotheses proposes that different large–scale landscape changes resulted in the fragmentation of species distributions that were formerly continuous and that this isolation promoted speciation. For example, it has been hypothesized that the uplift of the Talamanca and the Andean mountains resulted in the fragmentation of rainforests separated by drier biomes, and that speciation occurred in the absence of gene flow between isolated but ecologically similar demes (Grant *et al.*, 2006; Roberts *et al.*, 2006). Likewise, the refugia hypothesis states that diversification occurred in isolated demes as the result of forest fragmentation from cyclical climate oscillations during the late Pleistocene (Haffer, 1969; Haffer, 1997; Lynch & Duellman, 1997). An alternative and non-mutually exclusive hypothesis is that the principal effect of these landscapes changes on speciation was the formation of an ecologically structured landscape matrix upon which

subsequent diversification occurred (Smith *et al.*, 2014). In contrast to the allopatric view, this hypothesis emphasizes the role of natural selection in rain forest speciation. Under this landscape matrix model, diverging lineages are not necessarily geographically isolated, and speciation occurs as a by-product of climate-mediated divergent selection pressures along ecological gradients despite the presence –or absence– of gene flow (Endler, 1977; Schluter, 2009).

Under the vicariance models one may expect closely related species to show allopatric distributions and to have climatic niches more similar than those expected under a neutral model of evolution. While there is some evidence for this type of phylogenetic niche conservatism in tropical taxa, evidence has been mixed possibly because of the accumulation of ecological disparity once speciation is completed (Pyron & Wiens, 2013). On the other hand, ecological models that emphasized selection across ecological gradients (Lynch & Duellman, 1997; Graham *et al.*, 2004) predict that closely related species should have divergent ecological niches that are less similar than expected by a neutral model.

In this Chapter, I used a combination of geographic and climate data, multilocus molecular information and environmental niche modelling (ENMs) (Warren, 2012) to reconstruct the biogeographical history of the Harlequin poison frogs (*Oophaga* spp. Bauer, 1994). This genus represents a small group of colour polymorphic species distributed in a broad geographic extent, ranging from south-east Nicaragua in Central America to northern Ecuador in South America (Grant *et al.*, 2006) (Fig. 2-1).

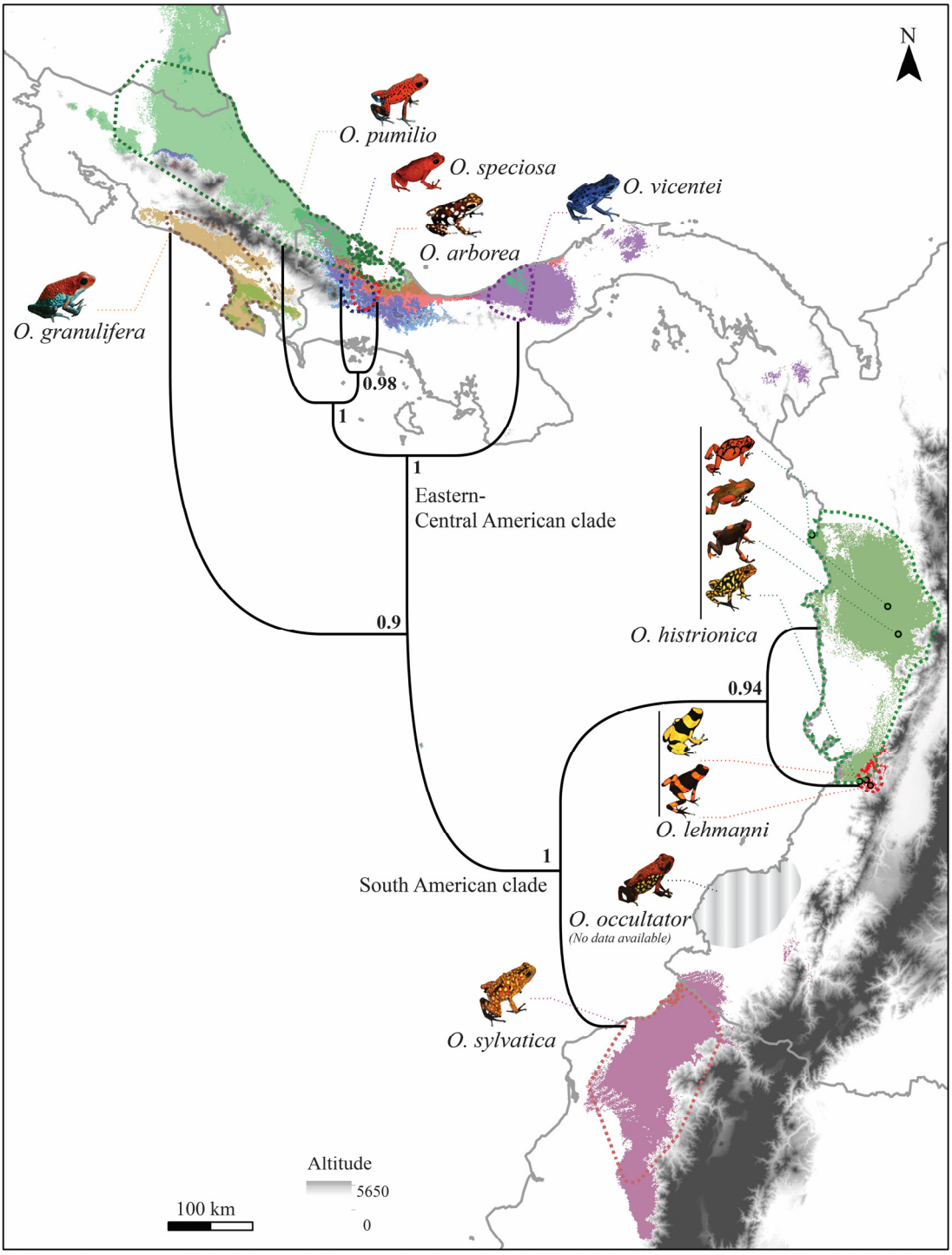


Figure 2-1. Map showing the distribution of *Oophaga* species. For each species, solid colours represent the suitable area as predicted by the ENMs while the areas delimited by the dotted lines represent their narrow geographic range defined as the geographic area covered by a convex Hull polygon including all known occurring localities. The maximum clade credibility (MCC) tree represents the phylogenetic relationships between *Oophaga* species as inferred in this Chapter (branch length is not scaled). Node numbers represent posterior probabilities (PP) as estimated in *BEAST. Altitude GIS layer is represented in grey colour.

Although there is some evidence that geographical barriers might play an important role in the diversification of some Central American lineages (Hagemann & Pröhl, 2007; Rudh *et al.*, 2007; Wang & Shaffer, 2008), the geographical proximity and overlapping distribution of some of the *Oophaga* species suggest that factors other than strict geographical isolation may also influence diversification (Maan & Cummings, 2009; Crothers *et al.*, 2011; Brusa *et al.*, 2013; Cummings & Crothers, 2013).

I used genetic data to generate a robust species tree based on mitochondrial and nuclear gene genealogies. Using this phylogenetic framework, I examined the role of geographic isolation and climatic niche divergence in the diversification of this group. In addition, because the Andean and Talamanca uplifts are likely to be one of the most relevant events in the history of this group, I examined the role of elevation gradients in *Oophaga* speciation. These data allowed me to test classical biogeographical hypotheses and to infer how geological processes have shaped the distribution and diversity of harlequin poison frogs.

2.2 Materials and Methods

2.2.1 Genealogical patterns and phylogenetic relationships among *Oophaga* frogs

The genetics dataset includes 394 published sequences from five mitochondrial genes (*12S, 16S, COI, CytB* and *tRNA-val*) and three nuclear loci (*SIAH1, H3* and *Rag1*), for all but one of the nominal *Oophaga* species (*O. arborea* Myers, Daly and Martínez, 1984; *O. granulifera* Taylor 1958*; O. histrionica* Myers & Daly, 1976; *O. lehmanni* Myers & Daly, 1976; *O. pumilio* Schmidt, 1857; *O. speciosa* Schmidt, 1857; *O. sylvatica* Funkhouser, 1956 and *O. vicentei* Jungfer, Weygoldt and Juraske*,* 1996). Because there are no data for *O. occultator,* I could not include this species in the analyses. Genbank accession numbers and the

number of sequences per species are presented in the Appendix S2A (Table S2A.1).

For each gene, aligned sequences (Clustal-W; Thompson *et al*., 1994) were used to estimate the best model of nucleotide substitution using JMODELTEST 2.1.4 (Posada, 2008) (Table S2A.2). Maximum likelihood (ML) and Bayesian (BI) gene genealogies were obtained using MEGA6 and BEAST 2.1.2, and species trees were estimated using Bayesian (*BEAST, (Drummond & Rambaut, 2007; Drummond *et al.*, 2012; Bouckaert *et al.*, 2014) and minimize deep coalescence (MDC) methods (Maddison & Knowles, 2006). Since gene genealogies were inferred using different datasets (*i.e.* individuals), I considered each locus as an independent partition. Alternative partitioning schemes considering codon positions (Table S2A.2) yielded identical species trees (Fig. S2B.1). Assumptions of neutrality and lack of intra-locus recombination were tested by evaluating the statistical significance of Tajima's D (Tajima, 1989) and the difference of sums of squares (DSS, TOPALI 2, (Milne *et al.*, 2009) respectively (See Appendix S2C).

2.2.2 Occurrence records and bioclimatic niche modelling (ENM).

A total of 313 unique presence records were obtained from my own field surveys, the IUCN Red List of Threatened Species (IUCN, 2012) and the Global Biodiversity Information Facility (http://www.gbif.org/species (Table S2A.3) after records from well outside the known geographical ranges of species were removed. Two species (*O. speciosa* and *O. arborea*) showed extremely restricted distributions (< 20 km^2) and I had a small number of records of these species for niche modelling (n < 10). I consequently increased the sampling size by including up to 40 random localities within their known spatial geographic range (IUCN, 2012), an approach similar to the systematic sampling described in Fourcade *et al.* (2014). The validity of this approach was further assessed by comparing the results generated from such

sampling to the niche model generated by the same number of field observations using *O. lehmanni* as a test case. I chose this species because I had highly reliable presence records and because it has a distribution similar in extension to that of *O. speciosa* and *O. arborea*. MAXENT-based niche modelling (ENM) (Phillips *et al.*, 2006) was performed using altitude and high resolution data for 19 bioclimatic variables (Digital Elevation Data 4.0, http://srtm.csi.cgiar.org/; WorldClim 1.4, http://www.worldclim.org. (Hijmans *et al.*, 2005) (Table S2A.4) with parameters and species-specific model tuning (Merow *et al.*, 2013; Radosavljevic & Anderson, 2014) as described in the Appendix S2C.

2.2.3 Geographic correlates of diversification: range overlap and niche filling.

I tested for patterns of phylogenetic relatedness in both geographic range and niche overlap by performing age-range correlations (ARC) (Berlocher & Feder, 2002; Fitzpatrick & Turelli, 2006) and principal coordinates of neighbour matrices (PCNM-RDA) between genetic distance and geographic overlap. Additionally, I estimated the degree of each species' "niche filling" (Nakazato *et al.*, 2010) in geographic (G-) and environmental (E-) space as the proportion of the space predicted to suit the species' ecological requirements that is actually known to be occupied by each species (Soberón & Peterson, 2011). Further details are presented in Appendix S2C.

2.2.4 Niche differentiation versus conservatism: sister-species comparisons.

To test whether sister-species' niches are more or less ecologically divergent than expected given the environmental backgrounds of the geographic regions where they occur, I used two conceptually different approaches. First, I calculated the niche overlap between species using the Schoener's D and Warren's I metrics in the ENMTOOLS package (Warren *et*

al., 2010) and used the background test (Warren *et al.*, 2008) to determine whether the estimated values were more similar than expected by chance. Second, I implemented the multivariate-based similar to that of method of McCormack *et al.* (2010) and for each principal component (PC), I compared the observed mean niche divergence with that of a bootstrap distribution (n=1,000 replicates) .

2.2.5 Niche differentiation versus conservatism: accumulation of disparity and history of climatic niche occupancy.

To evaluate the rate of overall bioclimatic niche evolution in relation to lineage diversification, I first evaluated niche occupancy disparity through time (DTT) (Harmon *et al.*, 2003). For each species and uncorrelated bioclimatic variables, I generated phylogenetic niche occupancy (PNO) profiles by combining the probability of occurrence in each grid cell with the corresponding MAXENT scores, binning the probability distribution into 300 evenly spaced categories (Evans *et al.*, 2009). Then, at each internal node of the *Oophaga* phylogeny, I calculated the relative niche occupancy disparity among and within subclades. Values around zero indicate that disparity is partitioned mostly among subclades, whereas values close to one indicate that disparity is high within subclades relative to total disparity across the entire phylogeny. The DTT curve was calculated using average squared Euclidean distances as implemented in the R-package 'geiger' 2.0.1 (Harmon *et al.*, 2008). The disparity index (MDI) was estimated over the first two-thirds of the DTT plots to avoid possible spurious disparity spikes close to the tip of the phylogeny (Harmon *et al.*, 2003). The MDI was used to test whether disparity within lineages is less than or greater than the median expectations of a neutral Brownian model of evolution (Harmon *et al.*, 2003; Evans *et al.*, 2009; Joly *et al.*, 2014).

To gain information on the directionality of niche evolution, I reconstructed the history of the evolution of climatic tolerances using a non-parametric approach and predicted niche occupancy density distributions (PNO) as implemented in the R 'phyloclim' package (Heibl & Calenge, 2013). For each species and non-correlated climatic variables, I randomly selected 500 values from the PNO distribution to estimate the maximum likelihood of ancestral climatic tolerance under the assumption of Brownian motion evolution (Schluter *et al.*, 1997; Joly *et al.*, 2014; Price *et al.*, 2014). Since this method relies on the reconstruction of potentially unrealistic ancestral states, I also used spatial evolutionary and ecological vicariance analyses (SEEVA) (Struwe *et al.*, 2011) to test if cladistics splits are associated with specific environmental splits and using Struwe's Ds (Struwe *et al.*, 2011) as a measure of ecological divergence between sister clades.

2.3 Results

2.3.1 Genealogical patterns and phylogenetic relationships among *Oophaga* frogs.

I found no evidence of recombination at any of the loci and Tajima's *D* values did not deviate significantly from the neutral expectation. The number of segregating sites, as estimated by *S* and *k*, differed by more than an order of magnitude between the two extreme values for *Rag1* and *SIAH1* (Table S2A.2). Uncorrected pair-wise genetic distances between lineages (*p*) ranged from ~ 1.7 -6%. For the set of mitochondrial genes *p*-distances were more pronounced, ranging for ~3-8%. The greatest genetic distances are those observed between *O. granulifera* and all other species (7-8% for mtDNA). Excluding this taxon, the maximum p-distance is ~ 3% (*O. histrionica – O. speciosa*). As a result, nearly all gene genealogies showed *O. granulifera* as a strongly supported (>90%) monophyletic clade (Fig. S2B.2). For all other

members of the genus *Oophaga*, the alleles in one species are as related to alleles in another species as they are to other co-specific alleles. Gene genealogies were therefore discordant, and did not show exclusive genealogical patterns. Bayesian and MDC species trees showed identical topologies and both resolved the phylogenetic relationship among *Oophaga* species with a strong node support (>90%). As shown in Figures 2.1 and 2a, *O. granulifera* represents the sister lineage to a group that comprises two distinct clades: all the other Central American species (*O. arborea, O. pumilio, O. speciosa* and *O. vicentei)* and all South American ones (*O. histrionica, O. lehmanni* and *O. sylvatica*) (Fig. 2-1).

2.3.2 Geographic correlates of diversification: range overlap and niche filling.

Oophaga poison frogs are distributed in a wide and disjunct geographic area where Central American species are separated from the South American species by the Isthmus of Darien (Fig. 2-1). Consequently, there is little overlap between distantly related species, however, sister taxa showed substantially overlapping geographic ranges (*O. histrionica – O. lehmanni:* 63% overlap; *O. speciosa* - *O. arborea*: 53% overlap). Accordingly, the geographic overlap decreased with increasing depth in the species tree (ARC r = −0.1147, $p = 0.006$), a trend expected if sympatric speciation has played a key role in the diversification of this group. However, the slope of the ARC-model is only marginally significant, and its intercept (0.3243, $p = 0.014$) does not go above 0.5 (as expected under a "pure" sympatric speciation model) (Fitzpatrick & Turelli, 2006). Randomization analyses also showed a negative relationship between genetic distance and geographic overlap (PCNM-RDA = -0.42, p = 0.032).

The degree of estimated niche filling on E- and G-space varied from 12.3% to 95.86% for all but one of the *Oophaga* species (Table S2A.5). Thus, the modelled niches served as accurate predictors of the species' geographic distribution (Fig. 2-1). For *O. histrionica,* the

niche predicted by ENM showed to be more restricted than its observed range in both E- and G-space, a result that is probably driven by the unclear taxonomic status of this nominal species (see Discussion).

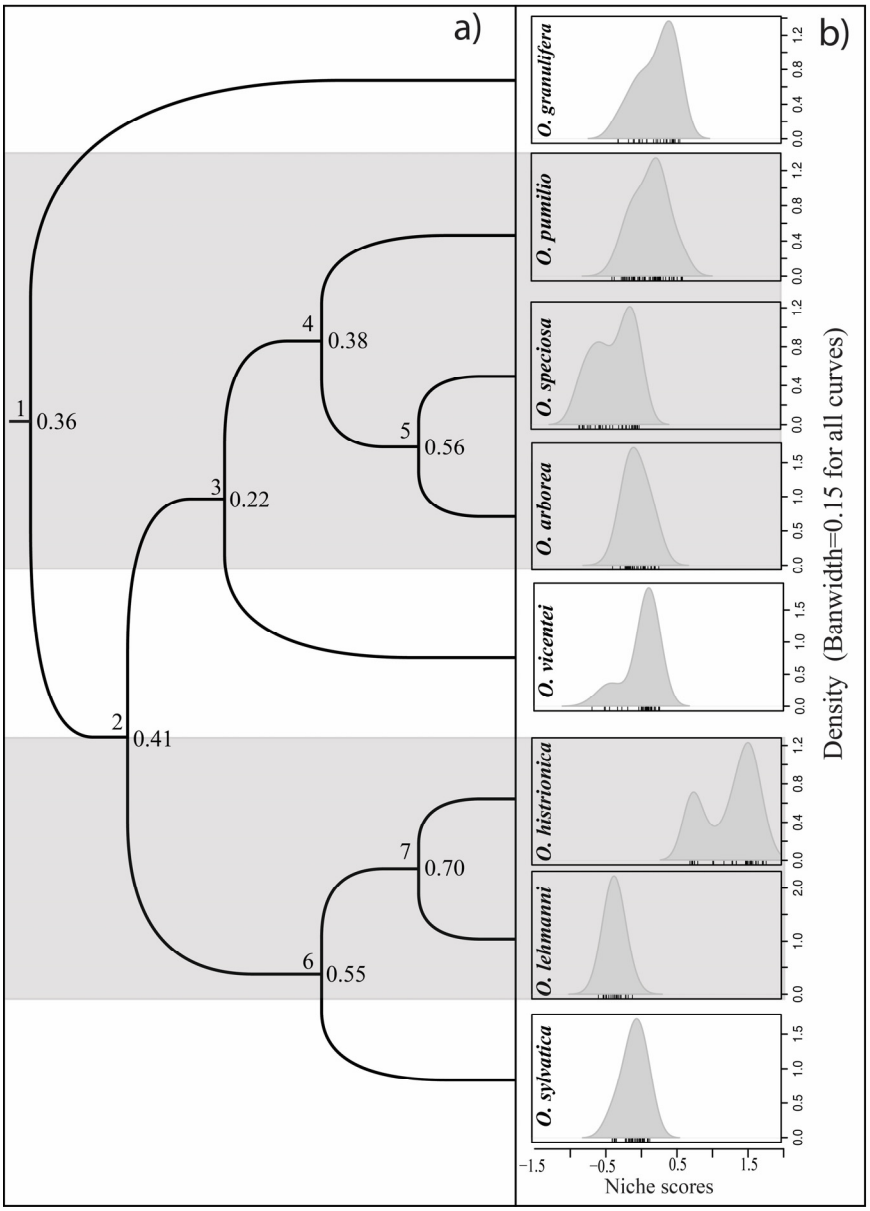


Figure 2-2. a). Maximum clade credibility (MCC) tree for *Oophaga* frogs. Values on each node represent node number (left) and divergence Struwe's Ds value (right). b) Kernel density estimation of climatic niche scores. The smoothing parameter (bandwidth) was selected using the Silverman's method (Silverman, 1986). Grey boxes represent species with overlapping geographic distributions.

According to the MAXENT models, three variables related to temperature oscillation explained most of the species distribution in Central-America. Whereas daily temperature oscillation (Bio2) is the most important single factor determining the distribution of *O. speciosa* (32%) and *O. pumilio* (24%), year round oscillation (combined Bio3 and Bio4) is most relevant factor in the other three species of this group (*O. granulifera,* 41%; *O. arborea*, 56%; *O. vicentei*, 38%). In contrast, species distributions in South America are mostly explained by precipitation variables. The total precipitation amount (Bio12) explains 58% of the *O. histrionica* distribution whereas the distribution of its sister species, *O. lehmanni*, is mostly determined by the amount of rainfall during the warmer months (Bio18, 38%). The amount and distribution of rainfall were weighted equally in the distribution of *O. sylvatica* (Bio12, 29%; Bio15, 27%; Bio18, 23%).

Values for measurement of niche overlap (Schoener's D) based on uncorrelated variables (Warren *et al.*, 2008) varied from 0.003 to 0.340. Pairwise comparisons between species with overlapping geographic ranges showed strong support for niche divergence in three out of four comparisons (Table 2-1). PCA analysis revealed four components (Eigenvalues >1) collectively explaining ~ 93% of the variance, with the first two vectors responsible for 52% and 23% of the variation. Variance along the first PCA axis is mostly explained by temperature-related variables and altitude (Table 2-2). Precipitation-related variables were mainly associated with the second and third niche axes (Table 2-2); however, they were correlated with latitude/longitude (Table 2-3). PCA results support the existence of ecological differences among species, and a high degree of niche overlap between the two main clades of the group (Fig. 2-3a). Multivariate analyses indicated significant niche divergence along PCA-axes in 59% of the pairwise comparisons (66/112). Divergence seems to be

particularly common among species with overlapping geographic ranges. In these species, 75% (12/16) of the pairwise comparisons rejected the hypothesis that species are distributed in identical climate space.

Table 2-1. Values for measurement of niche overlap: Schoener's D (lower diagonal) and Warren's I (upper diagonal). Values in bold indicate support for niche divergence while values in normal type indicate support for niche conservatism. Values within grey boxes indicate species pairs with overlapping geographic ranges (ns= no significant, p>0.05)

	O. vicentei	***O. arborea***	***O. granulifera***	***O. histrionica***	***O. Lehmanni***	***O. pumilio***	***O. speciosa***	***O. sylvatica***
O. vicentei		**0.141**	**0.028**	0.009	0.003	0.333 ns	**0.098**	0.132
O. arborea	**0.153**		0.097 ns	0.168	0.044	**0.126**	**0.309**	0.212
O. granulifera	**0.032**	0.082 ns		0.056	**0.004**	0.113 ns	**0.051**	0.031 ns
O. histrionica	0.008	0.152	0.060		**0.088**	0.025	**0.011**	0.110
O. Lehmanni	0.002	0.042	**0.004**	**0.086**		**0.004**	**0.004**	0.085
O. pumilio	0.324 ns	**0.117**	0.115 ns	0.024	**0.003**		0.118	0.195
O. speciosa	**0.091**	**0.314**	**0.048**	**0.013**	**0.004**	0.124		0.06*
O. sylvatica	0.138	0.315	0.04 ns	0.118	0.090	0.182	0.075 ns	

2.3.3 Niche differentiation versus conservatism: sister species comparisons.

To evaluate the degree and nature of ecological niche differentiation among the closest species pairs, I focused on the two pairs of sister taxa present in the species tree: *O. histrionica – O. lehmanni* and *O. speciosa* - *O. arborea*. Regardless of the measure of similarity used (Shoener's *D*: $D_{O.his\text{-}O.leh} = 0.086$, $D_{O.spe\text{-}O.arb.} = 0.314$; I: $I_{O.his\text{-}O.leh} = 0.088$, $I_{O.spe\text{-}O.arb.} = 0.309$), background tests supported ecological differentiation between sister pairs ($p < 0.001$). These species pairs are also clearly separated in multivariate bioclimatic space and estimates of niche overlap based on independent niche axes also support strong niche displacement between sister taxa (Table 2-3).

Table 2-2. Niche axes and bioclimatic variable loadings as implemented in the multivariate method (McCormack et al., 2010) using species data from both presence records and 1,000 random background points.

Variable	PC1	PC2	PC3	PC4
Annual mean temperature (°C)	8.10	2.85	0.84	0.92
Mean diurnal range (max temp-min temp)(monthly average)(°C)	1.57	0.01	10.85	38.23
Isothermality (BIO1/BIO7) * 100	0.13	11.05	2.81	16.49
Temperature Seasonality (Coefficient of Variation)	0.02	14.40	5.11	6.18
Max Temperature of Warmest Period (°C)	6.54	5.56	0.12	4.41
Min Temperature of Coldest Period (°C)	8.38	1.44	2.83	0.02
Temperature Annual Range (BIO5-BIO6) (°C)	1.73	3.66	18.00	13.63
Mean Temperature of Wettest Quarter (°C)	8.01	2.74	0.91	0.24
Mean Temperature of Driest Quarter (°C)	8.06	2.66	0.65	1.97
Mean Temperature of Warmest Quarter (°C)	7.72	4.05	0.41	0.58
Mean Temperature of Coldest Quarter (°C)	8.26	2.20	1.11	1.28
Annual Precipitation (mm)	5.56	5.30	6.44	0.88
Precipitation of Wettest Period (mm)	4.40	0.74	19.94	1.57
Precipitation of Driest Period (mm)	4.11	11.56	0.01	0.08
Precipitation Seasonality (Coefficient of Variation)	1.32	8.34	8.12	2.68
Precipitation of Wettest Quarter (mm)	4.87	1.59	16.46	2.19
Precipitation of Driest Quarter (mm)	4.43	10.96	0.00	0.03
Precipitation of Warmest Quarter (mm)	3.35	5.19	2.17	6.68
Precipitation of Coldest Quarter (mm)	5.37	2.65	2.89	1.86
Altitude (m)	8.05	3.06	0.32	0.08
Eigenvalue	**10.28**	**4.62**	**2.27**	**1.49**
Explained variance	**51.41%**	**23.08%**	**11.36%**	**7.45%**

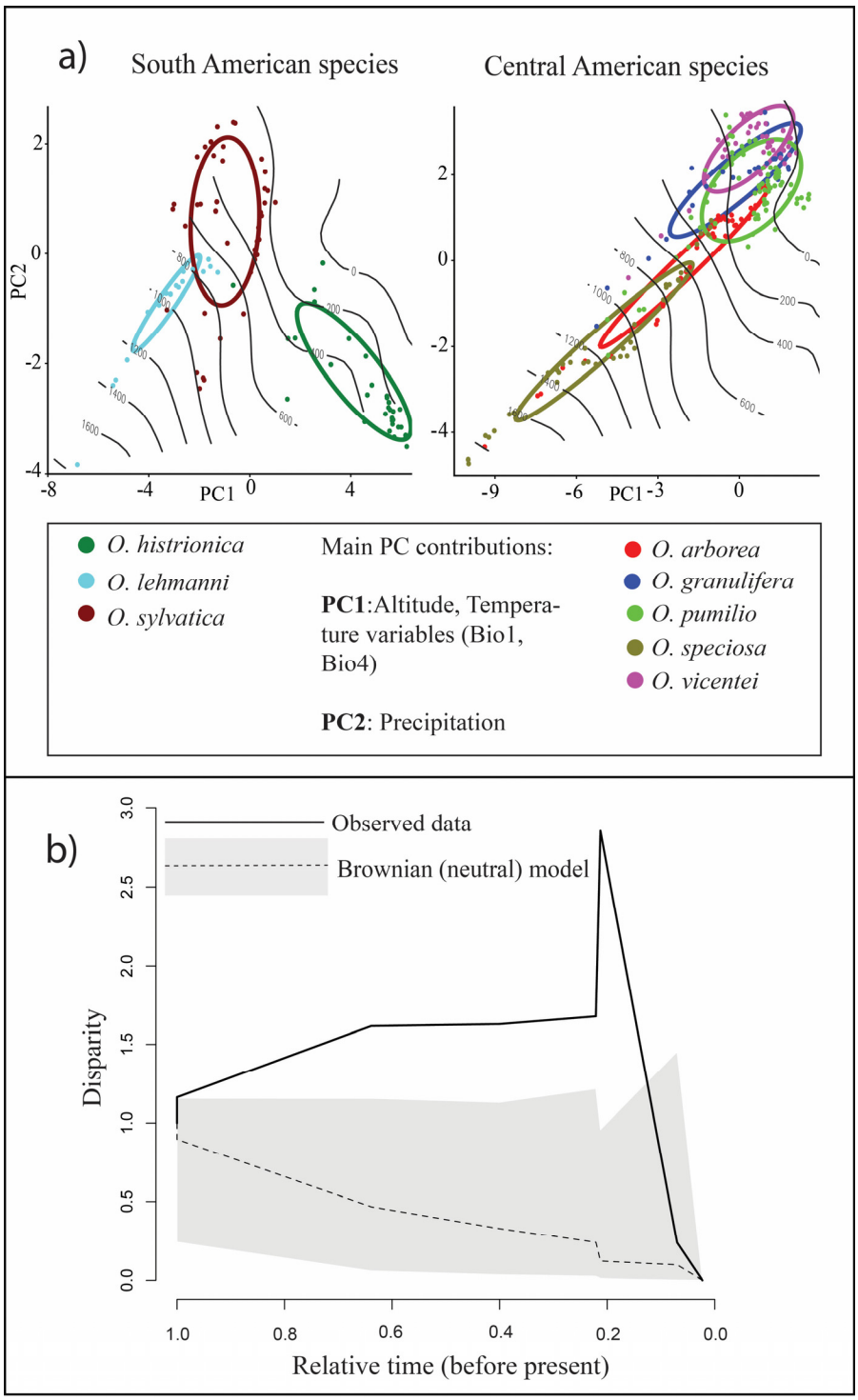


Figure 2-3. a) Two-dimension plots of the *Oophaga* species' multivariate niches based on the first two principal components for species within the two main clades. Each plot includes a representation of the altitude and indicates an altitudinal gradient replacement in sister species. Colour ellipses include 95% of each species' observations. b) Disparity through time (DTT) plot for *Oophaga* species' climatic niches (solid black line) and data simulated under a Brownian (neutral) model of niche evolution (dotted line). The grey area represents the 95% confidence interval of the Brownian model distribution.

Table 2-3. Niche divergence estimates based on independent multivariate niche axes for sister species and species with overlapping geographic ranges. Values in bold indicate support for niche divergence while values in normal type indicate support for niche conservatism. Additionally, all niche overlap values differed significantly between species pair after Bonferroni correction (P=0.002) (ns= no significant, p>0.05).

Pairwise comparison	**Niche axes**			
	PC1	**PC2**	**PC3**	**PC4**
O. pumilio vs. *O. arborea*	**0.671**	0.611 ns	**0.300**	0.035
O. pumilio vs. *O. speciosa*	**1.466**	**0.824**	**0.085**	**0.782**
O. arborea vs. *O. speciosa (sister taxa)*	**0.795**	**0.213**	0.215	**0.747**
O. vicentei vs. O. speciosa	**1.431**	**1.348**	**0.751**	**0.082**
O. histrionica vs. *O. lehmanni (sister taxa)*	**2.218**	**1.576**	**1.400**	1.536 ns
Variable contributions (biological explanation)	Temperature	Precipitation Elevation	Precipitation	Seasonality (Temp.-Precip.)
Correlation latitude	0.18	-0.57	-0.45	0.26
Correlation longitude	0.01	0.29	0.35	-0.09

2.3.4 Niche differentiation versus conservatism: accumulation of disparity and history of climatic niche occupancy.

Accumulation of overall ecological disparity analysed through DTT plots indicates a departure of observed patterns from the Brownian model of evolution (Fig. 2-3b). At most points in the phylogeny, the proportion of total disparity within clades was higher than expected by a Brownian motion model and is reflected in the positive MDI (0.79) (Fig. 2-3b). Ecological disparity shows a slow relatively stable progression of disparity at earlier times, and tends to accumulate reaching values above 2 at later stages in the evolutionary history of the group.

Table 2-4. Struwe's Ds values estimated by SEEVA (Struwe *et al*., 2011) for each phylogenetic split in the species tree. Ds is a measure from 0 to 1 on how skewed a distribution is between two different clades. Node numbers correspond to that in Fig. 2-2a. Bold values were significant after Bonferroni correction (P=0.0075).

Node	Temperature variables				Rainfall variables			Elevation	Mean
	Annual mean temperature	**Mean diurnal range**	**Isothermality**	**Temperature Seasonality**	**Annual Precipitation**	**Precipitation Seasonality**	**Precipitation of Warmest Quarter**		
1	0.06	**0.69**	**0.31**	**0.55**	**0.29**	**0.68**	**0.26**	0.07	0.36
2 (Central-South American clades)	**0.07**	**0.29**	**0.84**	**0.66**	**0.59**	**0.21**	**0.60**	0.01	0.41
3	**0.34**	0.04	**0.20**	**0.38**	**0.20**	**0.23**	**0.09**	**0.29**	0.22
4	**0.66**	**0.35**	**0.11**	**0.47**	**0.17**	**0.51**	**0.27**	**0.52**	0.38
5 (O. speciosa - O. arborea	**0.86**	**1.00**	**0.67**	**0.68**	0.10	0.05	**0.32**	**0.81**	0.56
6	**0.50**	**0.80**	**0.73**	**0.53**	**0.71**	**0.55**	**0.19**	**0.41**	0.55
7 (*O. histrionica - O. lehmanni*)	**0.96**	**0.43**	0.00	**0.73**	**0.93**	**0.66**	**0.91**	**1.00**	0.70

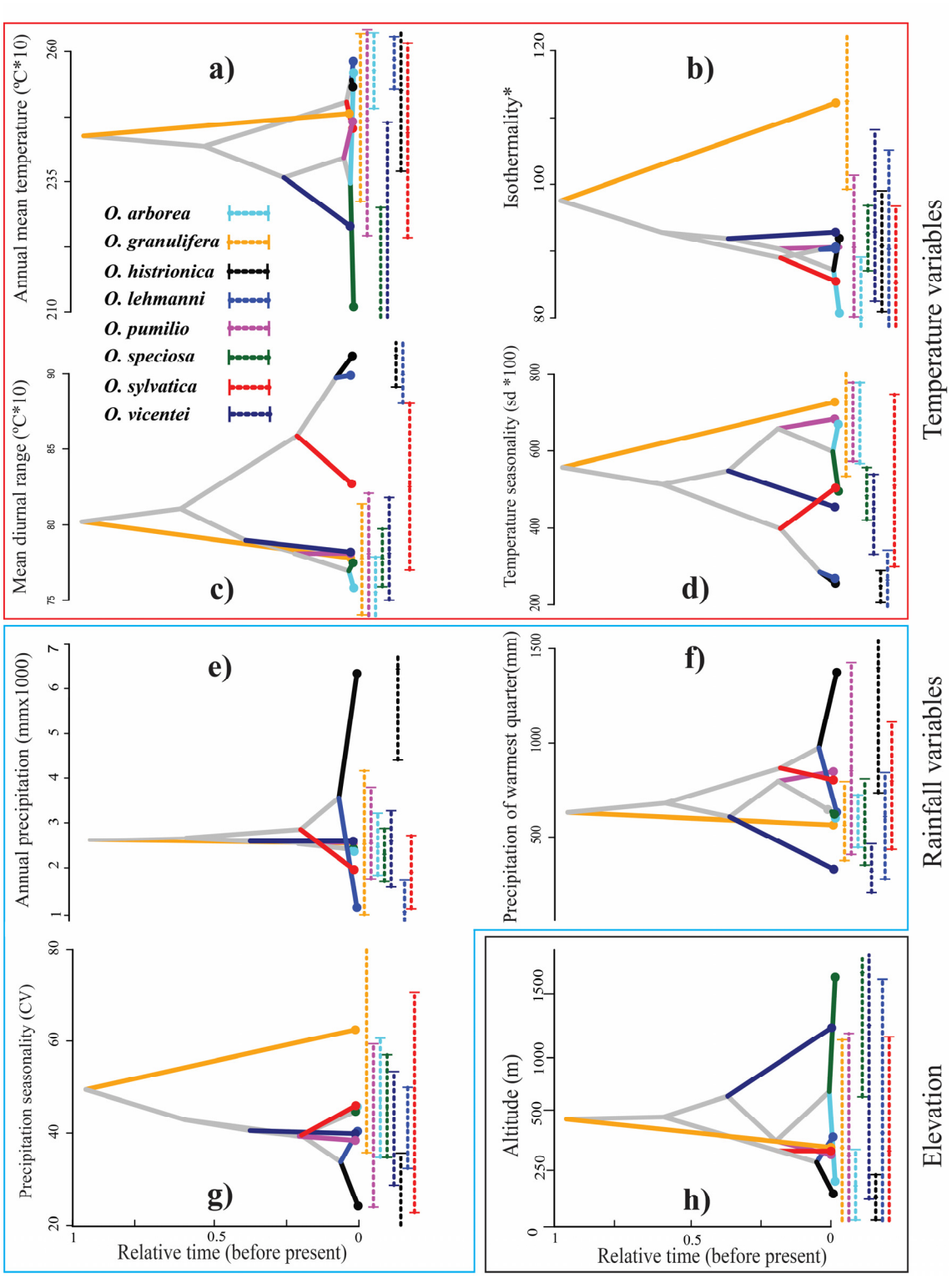


Figure 2-4. Inferred history of the evolution of climatic tolerances in *Oophaga* species (PNO profiles) for 8 non-correlated bioclimatic variables (a-h). The phylogeny corresponds to that of the maximum clade credibility (MCC) tree (Fig. 2-1 and 2-2). The 90% central density of climatic tolerances for each taxon is represented by the vertical dotted line and the mean value is represented by the end of each branch tip.

The history of climatic tolerances as inferred by PNO profiles showed a heterogeneous pattern of divergence on E-space (Fig. 2-4 a-h), suggesting a complex divergence in climatic preferences. South American species showed large shifts in the tolerance to different precipitation related variables. Particularly drastic is the case of annual rainfall (Fig. 2-4e). While *O. histrionica* inhabits very humid locations (5,000 to 8,000 mm year^{-1}), its sister species *O. lehmanni* occupies much drier zones with less than 2500 mm year^{-1} of precipitation. Central American species, on the other hand, showed very little variation along the climatic history for annual precipitation and all of them converge to a climatic requirement of 3000-3250 mm/year. Rainfall seasonality showed a similar trend of niche conservatism among Central American species but strong divergence within the South American clade (Fig. 2-4g). Precipitation requirements during the warmest quarter (Fig. 2-4f) also showed large shifts throughout the history of the group with most taxa converging to similar relative low precipitation levels (600-1000 mm) and the two species with the most southern distribution (*O. histrionica* and *O. sylvatica*) requiring more humid summers (>1000 mm). All *Oophaga* species are found in regions with very limited temperature oscillations (diurnal range: 7.6-9.5 ^{0}C; year-round range: 21-27 ^{0}C). As a result, isothermality (Bio3) values had remained relatively low (80-110) and fairly constant over time (Fig. 2-4b). Despite isothermality conservatism, tolerances for both diurnal and seasonal temperature oscillations showed a clear pattern of divergence across the entire phylogeny (Fig. 2-4c & 2-4d). Preference for high elevations (> 800m) have arisen independently in Central and South American clades resulting in strong divergence in altitude ranges between sister species ($Alt\text{-}_{O.his}$= 210 m *vs.* $Alt\text{-}_{O.leh}$= 748 m; $Alt\text{-}_{O.spe}$= 1002 m *vs*. $Alt\text{-}_{O.vic}$= 248 m; Fig. 2-4h).

SEEVA analyses showed a similar history of ecological divergence among *Oophaga*

frogs (Table 2-4). Tests on the seven phylogenetic nodes for each of the eight non-correlated environmental variables showed significant ecological shifts in 87% (49/56; 7 nodes, 8 bioclimatic variables) of the possible comparisons. In general, all ecological features showed substantial levels of divergence when average across nodes, however, basal splits tend to be associated with lower Ds values than those observed between recently diverged lineages (Table 2-4, Fig. 2-3a); a similar pattern to that inferred by the PNO profiles.

2.4 Discussion

2.4.1 Gene and species trees: A new phylogenetic hypothesis for *Oophaga*

The Neotropical poison frog genus *Oophaga* comprises nine nominal species that collectively have a wide geographic distribution and remarkable variation both within- and among-species in ecologically important traits, including aposematic colouration. To date, four studies have addressed the relationships among *Oophaga* frogs, but all reached different contradictory results (Clough & Summers, 2000; Grant *et al.*, 2006; Santos *et al.*, 2009; Brown *et al.*, 2011). A major source of uncertainty in these studies is that three of them failed to include more than one representative sample for most of the species. Accordingly, these studies did not have the power to distinguish between different hypotheses. In addition, all four used only mtDNA gene genealogies, which are less likely to represent the phylogeny of a group as accurately as trees based on multiple independent loci (Maddison & Knowles, 2006; Heled & Drummond, 2010). Here, I have combined extensive sampling and multispecies coalescent phylogenetic methods to provide a robust, alternative, phylogenetic hypothesis to the mtDNA trees previously published. The proposed phylogeny is significantly different from previously published mtDNA trees. Although this study concurs with that of Grant *et al*. (2006) regarding

the existence of a South and an eastern Central American clade, I found important differences in the structure and positions of species within each clade. One important difference between these trees is the location of *O. arborea*. While the mtDNA tree clusters this species with *O. granulifera*, the tree positioned *O. arborea* as sister to *O. speciosa* (99% ML, 0.98 PP) within the Central American clade including *O. pumilio*, *O. speciosa*, *O. arborea* and *O. vicentei*. Another significant difference is the position of *O. lehmanni* that I consider to be sister to *O. histrionica* (98% ML, 0.94 PP), a result consistent with two previous studies that consider *O. lehmanni* as part of the *O. histrionica* species complex (Lötters, 1992; Lötters *et al.*, 1999; Medina *et al.*, 2013).

Particularly interesting is the position of *O. granulifera,* which all studies (including this one) showed as a long-branch taxon distantly related to all other *Oophaga* species. Although this study does not include any outgroups, all existing rooted mitochondrial trees position *O. granulifera* as the basal taxon to all other *Oophaga* species, a pattern that is consistent with an old separation event. Assuming a rate of divergence of ~2 % per million years for mtDNA (Macey *et al.*, 1998; Schulte Ii *et al.*, 1998; Crawford, 2003) this split in the tree occurred during the Pliocene (~4 Ma) possibly as a result of a vicariance event associated with the rapid initial uplift of the Talamanca Mountains of Costa Rica and Panama (~4.5 Ma) (Abratis & Wörner, 2001; Gräfe *et al.*, 2002; Walther, 2003).

The reciprocal monophyly of the South and the eastern Central American clades and the similar mitochondrial divergence rate within (~3%) and between these clades (~3.9%) suggests a single colonization of South America by Central American *Oophaga* species following the final closure of the Isthmus of Panama (~3 MA) (Knowlton & Weigt, 1998; Montes *et al.*, 2012) and a rapid diversification at the beginning of the Pleistocene (~2 Ma) coinciding with or

postdating the last uplifts of the Talamanca and Andean mountains (Dzierma *et al.*, 2011; Echeverri *et al.*, 2015).

2.4.2 Geographic correlates of linage diversification in the *Oophaga* genus.

The analysis of geographic distributions in a phylogenetic framework, even with its potential caveats, is a useful step to assess the role of geographic isolation during diversification events (Nakazato *et al.*, 2010; Anacker & Strauss, 2014; Joly *et al.*, 2014). If, as predicted by Lynch and Duellman (1997), allopatric speciation has been the dominant process in this group, I expected the slope of the age-range correlation to be positive (Berlocher & Feder, 2002; Fitzpatrick & Turelli, 2006). Instead, I found the geographic overlap between species significantly decreased with increasing depth in the species tree, a pattern that is consistent with a predominance of sympatric/parapatric speciation within Central American and South American clades. Accordingly, I found a significant geographic range overlap between sister species (Fig. 2-1). The central question, however, is whether this particular pattern of geographic overlap observed in young and sister species is the result of rapid range expansions and secondary contact after speciation; or contrastingly, if speciation has been driven mainly in the presence of gene flow due ecological factors. Although my results are not conclusive, two indirect lines of evidence support the latter scenario: first, under a model in which substantial geographic overlap is caused by rapid secondary contact, and reproductive isolation is build up mainly in allopatry, the pattern of geographical overlap should also be evident in deeply divergent clades and not only in sister and young taxa. Second, while high levels of variation in morphological traits important for speciation (i.e. colouration, advertisement and mating calls) and strong genetic divergence have been driven quite rapidly in this group (1000-10,000 years) (Summers *et al.*, 1997; Wang & Shaffer, 2008); comparative evidence suggests that on average,

the achievement of complete secondary sympatry in the tropics takes on the order of millions of years (Weir & Price, 2011). This argument is likely to be particularly relevant in the case of *Oophaga* poison frogs, given the limited dispersal capacity of amphibians (Pyron & Wiens, 2013) and the lower levels of vagility when compared with other taxa (a mark-recapture study in *O. pumilio* showed that 70% of the recaptured individuals were found less than 20m from the marking site) (Nowakowski *et al.*, 2013). Thus, it is likely that speciation with gene flow has played a key role in the diversification of these two *Oophaga* clades. However, this conclusion should be taken with caution given the small geographic ranges of overlapping species and the limitations of the age-range correlation approach.

2.4.3 Niche filling patterns: taxonomic implications.

I have estimated the degree of niche filling by comparing the actual and predicted niches of each the nominal species of the *Oophaga* genus. For *O. granulifera*, *O. sylvatica* and *O. pumilio* the predicted geographical occurrence matched their actual distributions (niche filling > 85% in both G- and E-space), strongly suggesting that for these taxa, the models captured most of the ecological determinants of the species ranges. For *O. lehmanni*, *O. arborea*, *O. speciosa* and *O. vicentei,* the predicted suitability ranges were much wider than their observed distributions when projected in G-space (12 - 39% niche filling) while the models were shown to behave fairly well when evaluated in E-space for *O. speciosa* and *O. vicentei* (61-91 %). Different non-mutually exclusive factors are likely to account for such discrepancy. Micro-environmental factors, species interactions and dispersal limitation are known to limit species ranges, resulting in poor niche filling values (e.g. Peterson *et al*, 2002). In addition, all these species are rare and their estimated suitable habitats are based on small sample sizes. Because small datasets are more sensitive to potential outliers, this may have

resulted in an overestimation of the predicted ranges and the low niche filling values.

More interesting is the case of *O. histrionica*. For this species, the predicted suitable range was smaller than its known geographic distribution, a result that is probably driven by its unclear taxonomic status. As I mentioned above, this nominal species is likely to represent a complex of divergent allopatric lineages adapted to different bioclimatic conditions. A recent limited survey of genetic variation suggests the existence of at least two distinctive lineages occupying the northern and southern portion of the *O. histrionica* geographic range (Medina *et al.*, 2013). The ENM of *O. histrionica* shows a disjunct distribution overlapping with that of the northern and southern mitochondrial clades. This result strongly suggests that these two lineages are geographically isolated and ecologically distinct. Accordingly, the multivariate *O. histrionica*'s niche scores did not show a unimodal probability density function, but a bimodal one where each peak corresponds to each of the mitochondrial lineages (Fig. 2-2b). Although this analysis did not aim to clarify the taxonomic status of *O. histrionica,* these findings allowed me to propose the hypotheses that these lineages indeed represent two distinct (possibly sibling) species (Lötters *et al.*, 1999; Medina *et al.*, 2013).

2.4.4 Patterns of climatic tolerance among *Oophaga* species: niche differentiation versus conservatism.

A growing number of phylogenetic studies support the view that ecological speciation along elevation gradients are a key factor in the diversification of a wide range of taxonomic groups from plants to vertebrates (reviewed by Lawson &Weir, 2014). In this last part of the discussion, I present several lines of evidence to support this mechanism for diversification of *Oophaga*. If lineage diversification in these frogs has been facilitated by climatic niche

conservatism, I would expect support for allopatric speciation and strong climatic niche similarities between closely related species. Overall, my analyses revealed a very different pattern. First, niche models showed that the South and Central American clades of this genus have convergently evolved to similar patterns of geographic distribution and niche occupancy. While the species inhabiting higher (and drier) elevations tend to have small geographic ranges, the species inhabiting the lowland coastal margins tend to be more widely distributed. Second, geographic distribution and phylogenetic analyses showed that the parapatric species replacing each other along elevation gradients are indeed sister taxa, as predicted by the models of divergence along continuous ecological gradients. Accordingly, the results of the range-correlation analysis are consistent with parapatric speciation playing a key role in the diversification of this group. Third, there is strong climatic niche differentiation among sister taxa. For example, while *O. histrionica* inhabits places with high levels of overall rainfall (>6500 mm $year^{-1}$), its sister species, *O. lehmanni,* is confined to significantly drier zones (<2500 mm $year^{-1}$); an identical pattern to that found in sister Central American species. Fourth, PNO profiles and disparity analyses rejected neutral models of climatic niche evolution and showed large shifts in climatic niches throughout the history of the group. Thus, overall, the data presented here strongly suggest that speciation along climatic gradients (via niche divergence) has been a major evolutionary force behind the diversification of *Oophag*a poison dart frogs.

CHAPTER 3

ECOLOGY, MOLECULES AND COLOUR: DIVERGING LINEAGES AND CONSERVATION OF HARLEQUIN POISON FROGS *

IN MEMORIAM

"*I propose that each species has a distinctive life history, which include a series of stages that correspond to some of the named species concepts"*

Richard G. Harrison
1945-2016

3.1 Introduction

The delimitation of species boundaries is often ambiguous and we continuously debate the taxonomic status of many threatened organisms. In many cases, conflicting species concepts based on operational properties (e.g. reproductive isolation, Mayr 1942; ecological distinctiveness, Van Valen 1976; morphological diagnosability, Helbig *et al.* 2002; genealogical exclusivity, Baum & Donoghue 1995; etc) have led to confusing situations in which particular populations are considered to be species –or not– under alternative species concepts. For instance, the Florida panther, thought to be a unique (sub)species based on its morphological and ecological distinctiveness (Hedrick, 1995) is genetically considered an allopatric population of the North American cougar (*Puma c. cougar;* Culver *et al.* 2000).

While the adoption by many taxonomists of the General Lineage Concept (GLC) (de Queiroz, 2005b, a) has helped to ameliorate some taxonomic controversies, the delimitation of species is still far from trivial. Distinguishing between local adaptation, isolation by distance and species boundaries can be difficult in highly mobile organisms (Brown *et al.*, 2007). Similarly, inter- and intra-specific variation can also be very difficult to differentiate in those taxa characterized by allopatric lineages or low levels of morphological diversity (Barley *et al.*, 2013).

*This chapter has been submitted for publication. A preliminary version can be accessed at http://dx.doi.org/10.1101/050922

As a result, in the last decade we have witnessed the rapid succession of methodological approaches for identifying species boundaries. Thanks to the advent of sophisticated genomics tools and increasing computational power, new methods have been developed to delimit species based exclusively on genetic data (Heled & Drummond, 2010; Leaché & Fujita, 2010; Huelsenbeck *et al.*, 2011; Grummer *et al.*, 2013; Leaché *et al.*, 2014; Ruane *et al.*, 2014; Zhang *et al.*, 2014; Yang, 2015). Similarly, powerful statistical approaches have been proposed to use gaps in morphological variation or ecological discontinuities as the criteria for species delimitation (e.g. Valcarcel & Vargas 2010; Zapata &Jimenez 2012). Despite the widespread use of the above approaches, it is arguable whether species delimitation should be based on just a single line of evidence (Carstens *et al.*, 2013). Instead, an integrative approach should be able to provide the best inferences about species boundaries (Padial & De la Riva, 2006; Padial & De La Riva, 2010; Padial *et al.*, 2010).

Integrative taxonomy (Dayrat, 2005), combines several lines of evidence to support or reject the hypothesis that populations are independently evolving lineages (*i.e.* species; Padial & De la Riva 2009). Under this framework, different lines of evidence can either disagree over the number of species, or agree over the number of species but disagree over the assignment of individuals into species. So, how much agreement must different taxonomic characters show to consider a population (or a group of populations) a separate species? A prevalent strategy has been deriving species hypotheses by consensus, congruence or cumulation (*sensu* Padial *et al*. 2010) of separately analysed datasets (e.g. Cardoso *et al.* 2009; Vieites *et al*. 2009 ; Dejaco 2012; Yeates *et al*. 2011). One limitation of these methods is that they cannot extract the common signal that may emerge from the simultaneous analysis of all taxonomic characters. A second limitation is that the resulting species boundaries are likely to be affected by order in

which different characters (criteria) are introduced in the analyses. These methodological shortcomings, common in many taxonomic studies, need to be addressed.

Here, I applied a multivariate approach to analyse phenotypic, environmental and genetic variation that can serve to assign independent evolutionary lineages and/or operational taxonomic units. For this purpose, I performed a comprehensive study of several Harlequin poison frog populations (*Oophaga histrionica* Myers and Daly 1975). I have two reasons to focus on this system. first, polytypic taxa (phenotypic variation over geographic localities) present important challenges for lineage delimitation. Second, specimens of these frogs are actively sought in the wildlife pet trade. Intense illegal trafficking and habitat loss associated with the proliferation of illegal crops have pushed these frogs to near extinction. Thus, it is critically important to delimit conservation units (at or below the species level) that can guide the implementation of the appropriate taxon-based conservation legislature (e.g. CITES trade treaty; Lynch and Arroyo 2009).

Harlequin poison frogs inhabit the lowland Pacific rainforests of the Colombian Chocó. Individuals from different populations can either be striped, spotted or relatively homogenous, and their colours range from bright red and yellow to dull green and blue (Fig. 3-1). This phenotypic variability combined with their patchy distribution and strong population structure suggests that *O. histrionica* may, in fact, be a complex of several different species. However, early taxonomic studies (Berthold, 1846; Myers & Daly, 1976) and preliminary genetic data from geographically scattered individuals (Grant *et al.*, 2006; Brown *et al.*, 2011) have failed to recognized the striking phenotypic diversity and potential ecological (environmental) divergence in this group (Lötters, 1992; Lötters *et al.*, 1999; Medina *et al.*, 2013; Posso-Terranova & Andrés, 2016a). To date, only two nominal species are recognized: *O. histrionica*

(*sensu lato*) and *O. lehmanni* (Myers & Daly, 1976). While the former taxon has a widespread distribution, *O. lehmanni* is a micro-endemic species restricted to a few small submontane wards of the central Chocó (Fig. 3-1a).

a)
Bahia Solano
El Valle
Atrato River
NW
Tutunendo
Barranquito
Pacurita
NE
La Victoria
Angostura
El Salero
Santa Cecilia
CENTRAL
San Juan River
SOUTH
Anchicaya
50km
La Choza, El Cocal
Cauchal
La Cruceta, El Filo
b)
c)
Bioclim. 1
Bioclim. 2
Bioclim. 3
Bioclim. 4
Bioclim. 5
Bioc. 6
d)
Clade I
Clade II
Clade III
Clade IV
O. lehmanni
e)
O. histrionica
O. lehmanni
O. lehmanni*
Lineage I
NW
Lineage II
NE
Lineage III
Central
Lineage IV
South

Figure 3-1. Left to right. a) Map of the sampling localities. b) Representative examples of phenotypic variation found in the "histrionica" complex and posterior assignment probabilities to each of the nine distinct morphological groups inferred by the DAPC analysis of 50 morphological variables (see Fig. 3-2). c) Schematic representation of the different bioclimatic niches occupied by different populations as inferred by the NMDS analysis of 19 bioclimatic variables and altitude. The multivariate representation of the Gaussian bioclimatic clusters is provided in Figure 3-3. d) Grey: Monophyletic mitochondrial clades (I to IV and *O. lehmanni*) as inferred by Bayesian and ML methods. A detailed topology of the phylogenetic tree including *O. sylvatica* and *O. occultator* as outgroups is shown in Figure 3-4. The corresponding parsimony haplotype network is available in Figure 3-5. Colour: Posterior probability of assignment to five genetic clusters based on Bayesian analysis of variation at microsatellite and transcriptome-derived haplotypes. Graphical representation of the corresponding PCoAs are provided in Figures 3-6a and b. e) In grey, species delimitation based on Myers and Daly (1976) and Funkhouser (1956). Coloured-boxes represent the lineage delimitation resulting from the multivariate integrative analyses of this study.

In this study, I used next generation sequencing (NGS) to obtain haplotypes that are then analysed using multivariate clustering approaches. This approach, similar to the species delimiting method of Hausdorf and Hennig (2010) and the species assignment method of Edwards and Knowles (2014), combines transcriptome polymorphisms with mtDNA and microsatellites variation to generate a comprehensive genetic dataset that is then combined with patterns of phenotypic and environmental divergence. The resulting common signal is used to delimit potential independent evolutionary lineages. To evaluate the power of our approach to identify these independent evolving lineages (i.e. species), I compared its results to those obtained by other "traditional" integrative methods (see Edwards and Knowles 2014). This analyses revealed the existence of at least five different lineages within the *O. histrionica* complex, some of them occurring in very small isolated populations outside any protected areas. More broadly, this study exemplifies how targeted amplicon sequencing and multivariate statistical analysis of different datasets can be integrated to unravel the complexities underlying phenotypic differentiation in polytypic taxa.

3.2 Materials and Methods

3.2.1 Specimen sampling

I sampled 310 individuals in 15 different geographic locations across the known geographic distribution of the *O. histrionica* complex (Myers & Daly, 1976; Lötters, 1992; Lötters *et al.*, 1999) (Fig. 3-1a). For each of the sampled individuals, I recorded its geographic location using a handheld GPS unit. High-resolution images and tissue samples were then collected as follows: raw digital pictures (Nikkon Electronic Format: NEF) were taken using a Nikkon D610 digital camera against a Spectralon grey reflectance standard (Labsphere,

Congleton, UK). Images were then linearized with respect to light intensity (Stevens *et al.*, 2007) and stored as TIFF images (lossless compression) to preserve fidelity for further analysis. Tissue samples were obtained by toe-clipping (Phillott *et al.*, 2011). Toe-clips were conserved in 95% ethanol until DNA extraction. Genomic DNA was extracted using the MasterPure^{TM} DNA purification kit (Epicentre, Madison, WI, USA) following the manufacturer's recommendations. All samples were obtained in compliance with the Ethic Committee of the Universidad Nacional de Colombia (sesión del 22 de Julio, Acta No.03, July 22nd 2015), under the research permits granted by Autoridad Nacional de Licencias Ambientales ANLA, Resolución 0255 del 12 de Marzo de 2014 and Resolución 1482 del 20 de Noviembre de 2015.

3.2.2 Morphological clustering

Independent evolutionary units could be delimited on the basis of gaps in patterns of morphological variation (Zapata & Jiménez, 2012). I estimated the number of distinct morphological clusters using hierarchical and discriminant grouping methods. Dorsal and ventral pictures of each specimen were scaled using ImageJ v. 1.48 (Sheffield, 2007) and a known metric reference. For each specimen, a total of 49 variables (n_{dorsal}= 25; $n_{ventral}$= 24) related to morphology, pattern of colouration and colouration intensity (Table S3A.1) were evaluated. All variables were standardized prior to subsequent analyses. First, I performed a hierarchical cluster analysis as implemented in PAST v. 3.08 with bootstrap- and PERMANOVA-based significance tests (Hammer *et al.*, 2001). In addition, we also performed a discriminant analysis of principal components (DAPC) as implemented in the R package *Adegenet* v 2.0.0. (Jombart *et al.*, 2010). This method has the advantage that allows for a probabilistic assignment of specimens to each group. To determine the optimal number of morphological clusters (K_M), we ran the sequential *K*-means (*find.clusters, K*=1-19) algorithm

and identified one showing the lowest Bayesian Information Criterion (BIC) (Jombart *et al.*, 2010). To avoid overfitting, the individual assignment of specimens was based on 11 axes (*optim.a.score*) representing 86% of the variation in the dataset.

3.2.3 Environmental clustering

To identify environmental discontinuities and determine the number of environmental clusters (K_E), I performed a non-metric multi-dimensional scaling (NMDS) analysis as proposed by Edwards and Knowles (2014). Briefly, for each sampled individual, we extracted high resolution (30 arc-seconds) environmental data from 19 bioclimatic variables (www.worldclim.org) along with altitude data (STRM90 http://srtm.csi.cgiar.org/) (Hijmans *et al.*, 2005). Bioclimatic data were standardized and reduced to PC-scores along 4 axes (explaining > 90% of the variation) using the *dudi.pca* function in the Vegan R-package (Oksanen *et al.*, 2013). Next, geographic distances among sampled specimens (function: *Distance*, Diva-Gis v. 7.5; http://www.diva-gis.org/) (Hijmans *et al.*, 2001) were converted to continuous variables (function: *pcnm*. Vegan R-package) and combined with the bioclimatic scores before generating the final Gower-distance environmental dataset. Then, I used non-metric multi-dimensional scaling (NMDS) and one-way PERMANOVA to estimate K_E.

3.2.4 Genetic clustering

Mitochondrial DNA (mtDNA): Mitochondrial genetic subdivision is often considered a leading criterion of interspecific differentiation. To determine the number of mitochondrial clusters (K_{MT}), I applied phylogenetic and multivariate procedures. For a subsample of 107 individuals (including *O. sylvatica* (n=6) and *O. occultator* (n=4) as "*bona fide*" control species), I amplified and Sanger-sequenced a 523 bp of the COI gene (Medina *et al.*, 2013).

Aligned sequences (*Clustal-W)* (Thompson *et al.*, 1994) were used to estimate the model of nucleotide substitution using the corrected Akaike Information Criterion as implemented in jModelTest v2.1.4 (Posada, 2008). MtDNA phylogenetic trees were constructed under different Bayesian and maximum likelihood optimality criteria and a TCS haplotype network was also plotted (Appendix S3B). Individual pairwise genetic distances were calculated using MEGA6 (Tamura *et al.*, 2013) and a Principal Coordinate analysis (PCoA) was carried out upon the similarity matrix (Reeves & Richards, 2011). Finally, to delimit independent evolving linages I also used a generalized mixed- Yule-coalescent (GMYC) method capable of fitting within- and between-species branching models (Zhang *et al.*, 2013) (Appendix S3B).

Microsatellites and transcriptome-derived haplotypes: Twelve microsatellite loci were isolated from a di-, tri- and tetra-mer enriched library constructed from a single *O. histrionica* specimen using pyrosequencing technology (454GS-FLX. Roche Applied Science) as described in Andres and Bogdanowicz (2011) (Table 3A.2). All individuals (n= 310, $n_{localities}$= 15) were genotyped using multiplex-PCR and fluorescently-labelled primers as described in Appendix S3B. To further identify polymorphic nuclear loci for multiplex PCR, a transcriptome-based approach was used. *De novo* skin transcriptome sequencing and assembly (Table S3A.3) from a single *O. lehmanni* individual was carried out as described in Appendix S3B. The resulting transcriptome was used to design PCR primers for 50 nuclear markers (250-350 bp), each of them representing a single randomly selected unigene. To reduce ascertainment bias, I selected eight individuals from a broad geographic range to screen each locus for robust PCR using genomic DNA as template. After stringent selection, 20 remaining loci were used to genotype all 310 specimens in two multiplex PCR groups using the QIAGEN Multiplex PCR Kit. Individual specimens and loci were barcoded using Illumina™ S5 and N7

Nextera primers following the manufacturer's conditions. Amplicons were paired-end sequenced (2X 250 bp) on a Illumina MiSeq at the Cornell University's BioResource Center.

Post-sequencing data processing required extracting reads from the Miseq run and assigning them to the appropriate specimen and locus with a custom *Perl* script that discarded low quality reads, trimmed adapters, overlapped paired-reads, identified reads corresponding to each locus, collapsed identical reads for each individual, and identified the two most frequent haplotypes for each individual at all loci (quality score: Q12, minimum overlap: 20, mismatch rate: 0.05). At any given locus, an individual was scored as heterozygous if the minor allele frequency was ≥ 20%. Individuals that fail to amplify in more than 80% of the loci were excluded from the dataset, and any sample by locus cells with fewer than 10 total reads were considered missing data.

The number of genetic clusters based on microsatellite loci (K_{SAT}), transcriptome-derived haplotypes (K_{TRN}), and both datasets combined (K_{NUC}) was estimated using three allele frequency-based methods (STRUCTURE, STRUCTURAMA, and PCoA). While the first two methods estimate the number of clusters by minimizing the deviations from Hardy-Weinberg equilibrium, the latter one finds clusters based on the Nei's distance (Pritchard *et al.*, 2000; Hubisz *et al.*, 2009; Peakall & Smouse, 2012). Detailed information on the different models, parameters and run conditions is described in Appendix S3B.

3.2.5 Total-evidence lineage delimitation

The multivariate delimitation approach described in this paper takes advantage of the GLC species concept, and is based on the conceptual and methodological frameworks of Hausdorf and Henning (2010) and Edwards and Knowles (2014). First, a provisional lineage hypothesis is formulated based on a small balanced geographical sampling of the focal system

and the assumption that independent lineages can be defined as distinguishable groups (i.e. multivariate clusters) of individuals with similar phenotypic, genotypic and climatic-environmental characteristics (see Discussion). Second, I weighted the strength of the evidence supporting the proposed lineages hypothesis by evaluating the ability to detect and assign new specimens to the provisional clusters. Because this multivariate approach does not necessarily take into account historical patterns of diversification, I compared the results to those obtained by using a Generalized Mixed Yule Coalescent (GMYC) model capable of fitting within- and between-species branching models (Zhang *et al.*, 2013).

Accordingly, for a first sample of individuals (n=86, 4-12 ind./population), we generated a new set of standardized orthogonal variables containing the same information as the original datasets using multivariate methods (morphology: PCA, environmental and genetic: PCoA). Then, distance matrices for each data-type were calculated (morphology and environmental: Gower-distances; mtDNA: TN93-distance; nuclear loci: Nei's distance) and standardized using NMDS (*isoMDS*, MASS R-package). To determine the provisional number of independent lineages, I followed Edwards and Knowles (2014) and retained four dimensions for each data-type and determined the number of Gaussian clusters present in the subsample using the R-package *mclust* and the Bayesian Information Criterion (Fraley & Raftery, 2002). Additionally, I also performed hierarchical clustering (average method) with 1000 bootstrap repetitions as implemented in PAST v. 3.08 (Hammer *et al.*, 2001). To assess the strength of the provisional lineages delimitation hypothesis, I conducted an identical series of analyses on an independent, much larger dataset (n= 214 individuals) and assessed the ability to recover the same clusters with high individual assignment probabilities. All stress values for NMDS analyses were below 10%, with an absolute maximum of 6%.

3.3 Results

3.3.1 Morphological and environmental clustering

The morphological hierarchical clustering applied over individuals of the *O. histrionica* complex revealed the presence of nine clearly defined morphological clusters with high node support (>82%; PERMANOVA Bonferroni-adjusted P<0.003; Fig. 3-1b and Fig. 3-2). The optimal number of morphologic clusters (K_M) as defined by DAPC (BIC) was also nine. As a result, both clustering methods were highly congruent in their individual assignments (Fig. 3-1a). Posterior assignment probabilities (PP) for K_M=9 indicated that 93% of individuals were correctly assigned to their morphologic clusters (PP>0.9); 4.7% were considered as ambiguously assigned (PP<0.9) and only 2.3% were incorrectly assigned (*i.e.* assigned to a different cluster with PP>0.9).

Gaussian clustering based on a NMDS analysis of bioclimatic distances partitioned the sampled geographic range of the *O. histrionica* complex into six distinct ecological clusters (K_E= 6, stress=0.04) with statistical significance after Bonferroni correction (P<0.0001; Fig. 3-1c): while *O. histrionica* (*sensu lato*) encompasses four potential bioclimatic niches, putative *O. lehmanni* specimens form two different clusters that correspond to two different morphological clusters (Fig. 3-1c and Fig. 3-3).

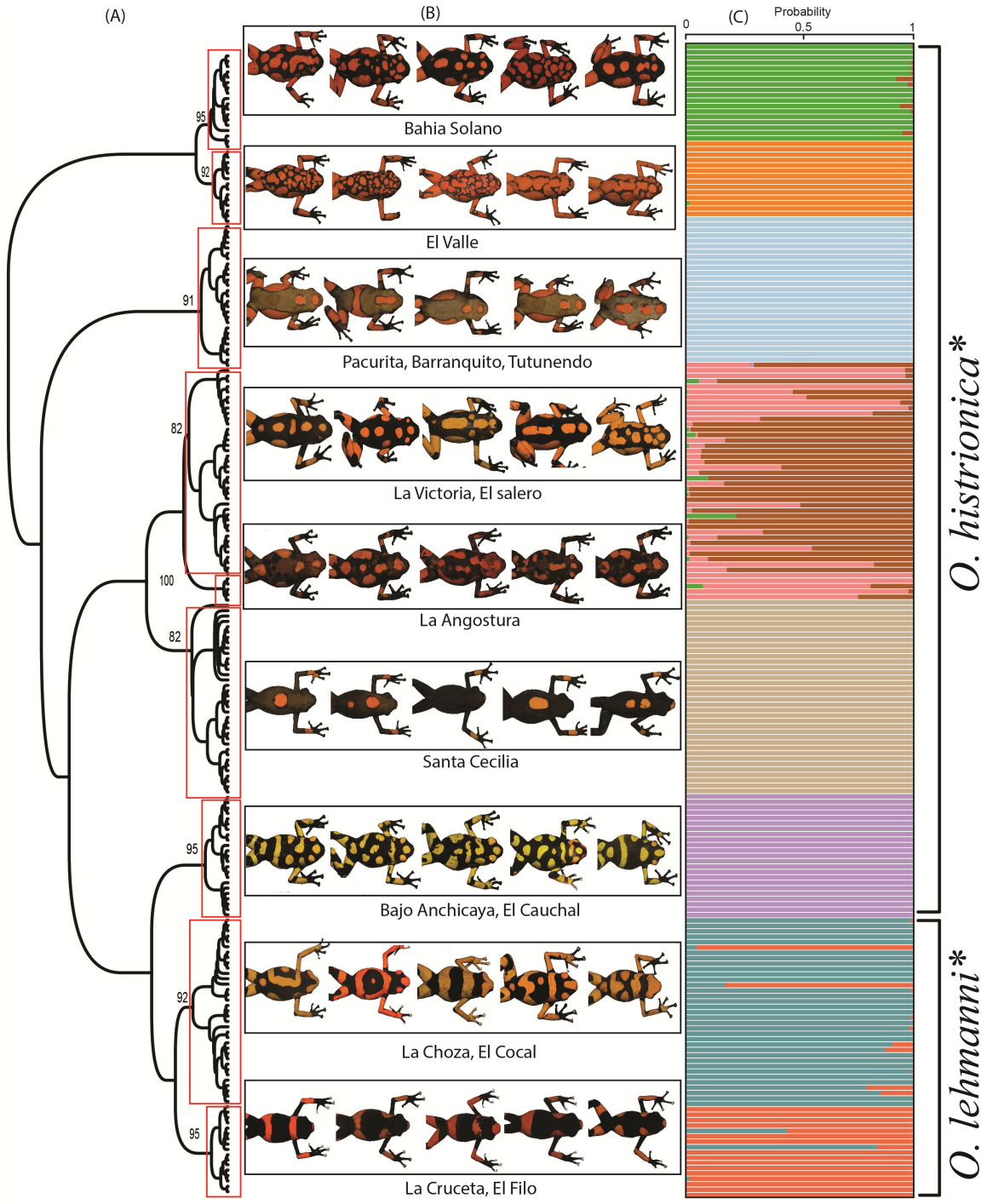


Figure 3-2. a) Morphological hierarchical clustering of individuals of the *O. histrionica* complex. Clusters with node support >82% are shown in red boxes. b) Individual morphologic variation within clusters and their respective sampling location. c) Posterior assignment probabilities (PP) for each individual based on DAPC analysis of 50 morphologic variables. *. Current taxonomic delimitation (Myers & Daly, 1976)

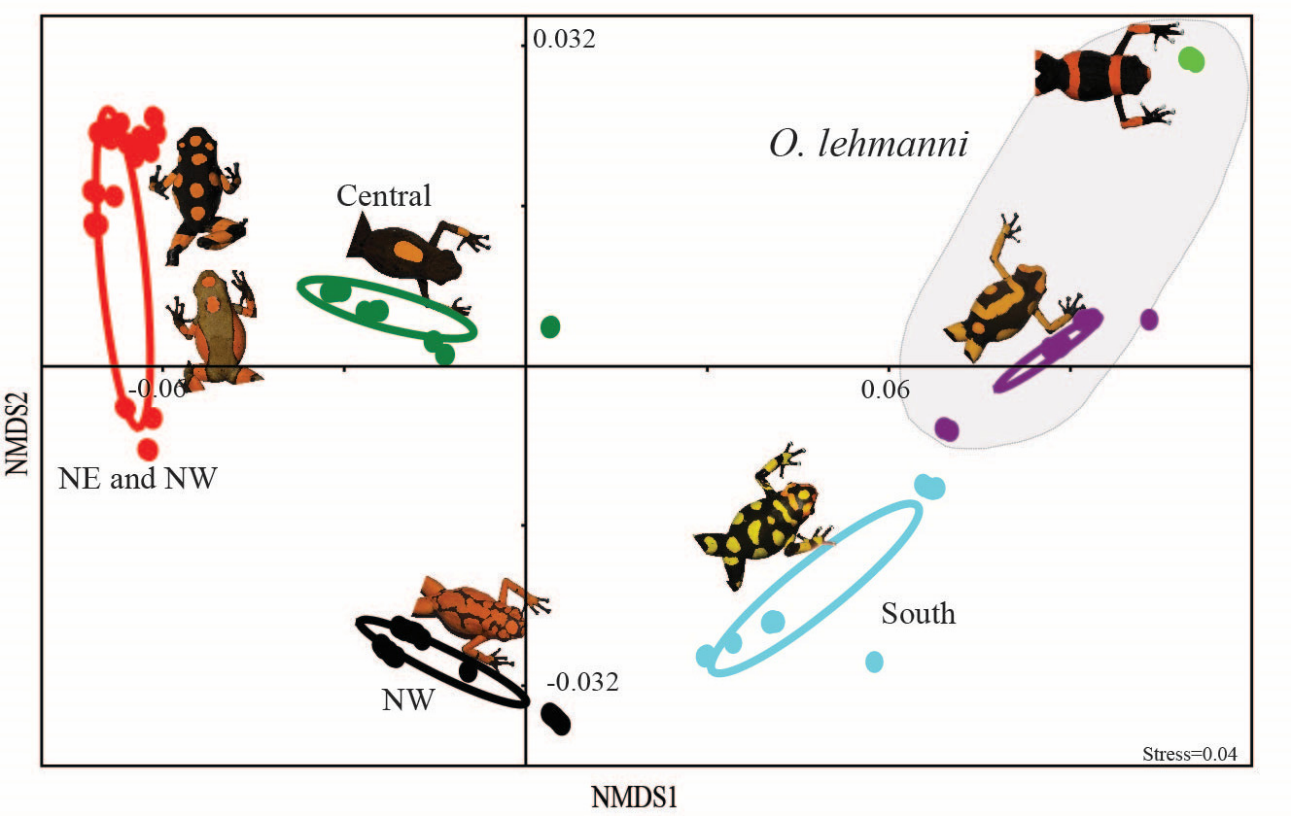


Figure 3-3. Environmental clusters by NMDS of climate data. Six potential climatic niches were found that correlates with geographic distribution. Clusters within coloured ellipses were significantly different (P<0.0001 after Bonferroni correction) and include 95% of the data. Clusters within the shadowed ellipse correspond to *O. lehmanni* populations.

3.3.2 Genetic clustering

Bayesian (BI) and Maximum Likelihood (ML) COI gene genealogies showed identical topologies and revealed the presence of seven distinct mtDNA clades representing *O. sylvatica*, *O. occultator*, *O. lehmanni* (defined by phenotype and sampling location), and 4 monophyletic, geographically structured clades (I-IV) within *O. histrionica* (*sensu lato*, Fig. 3-1d and Fig. 3-4): Clade I (NW-populations), Clade II (NE- populations), Clade III (Central populations), and Clade IV (South populations) –See also haplotype network provided in Figure 3-5 –. As expected, on average, the *p*-distances between the "*bona fide* species" (*O. occultator* and *O.*

sylvatica), and the *O. histrionica* complex clades were greater (range: 4-6%) than those observed among the four *histrionica* clades (range: 2.2%-2.9%). The single-locus delimitation analysis by GMYC strongly supported the hypothesis that each of these mtDNA clades represent an independent evolutionary lineage (PS>0.96 in all cases).

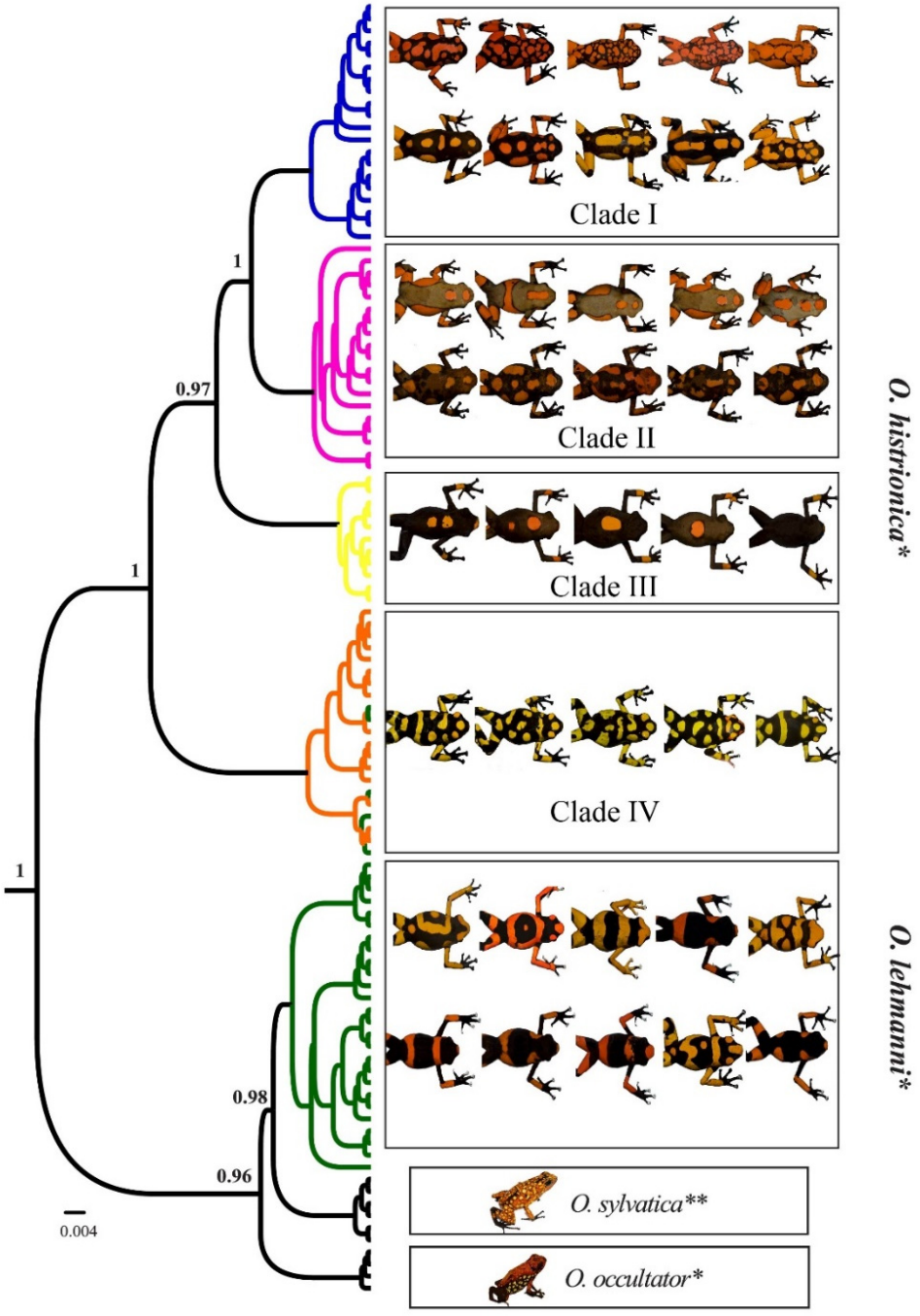


Figure 3-4. Mitochondrial lineage delimitation of the *O. histrionica* species complex. Numbers in nodes represent Bayesian posterior probabilities (speciation probability, SP) and only values higher than 0.95 are shown. The currently recognized *O. sylvatica* and *O. occultator* nominal species are included as control species. Species names in the figure correspond to the current taxonomic classification (*: *sensu lato* Myers and Daly 1976. **: Funkhouser 1956).

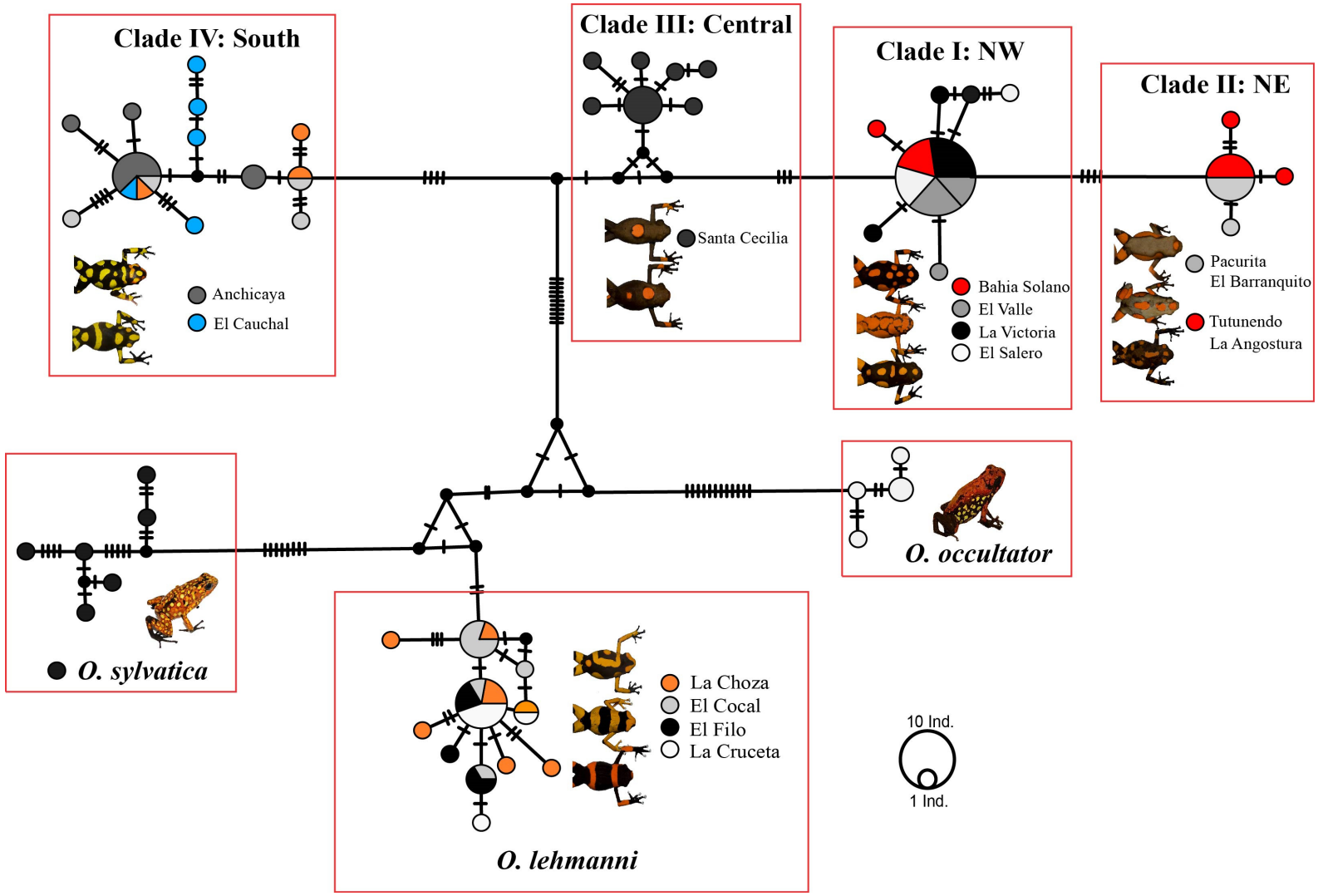


Figure 3-5. Mitochondrial (COI) parsimony haplotype network of individuals of the *O. histrionica* species complex. Clade I-IV correspond to those mitochondrial clades of Figure 3-4. *O. sylvatica* and *O. occultator* are included as control species.

Clustering analyses of the *O. histrionica* complex based on the two nuclear datasets yield consistent partitions that, in turn, were very similar to that based on mtDNA. Although the PCoA results were mostly concordant, the microsatellite dataset supported a five-cluster partition (K_{SAT}= 5, Fig. 3-6a) while transcriptome-derived dataset supported six clusters (K_{TRN}= 6, Fig. 3-6b). The source of discordance is the adscription of the individuals from the localities of La Choza and El Cocal. The analysis of the transcriptome-derived haplotypes suggested that these populations represent an intermediate cluster between the Southern populations of *O. histrionica* (*sensu lato*) -Anchicaya and El Cauchal- and the *O. lehmanni* populations of La Cruceta and El Filo (Fig. 3-1a). In contrast, the microsatellite-based analysis clustered these individuals with either the Southern *O. histrionica* or the *O. lehmanni* cluster. These results are consistent with La Choza and El Cocal being admixed populations. As expected under this admixture hypothesis, La Choza-Cocal specimens have intermediate phenotypes between those of *O. lehmanni* and Southern *O. histrionica* (Fig. 3-1b), and either Clade IV or *O. lehmanni* mtDNA haplotypes (Fig. 3-1d).

Bayesian clustering (STRUCTURE) analyses for the combined data set (microsatellites + transcriptome haplotypes) showed that the estimated log probability [*Ln P(D)*] plateaued at K_{NUC} = 6 with the second-order rate of change of *K* (*ΔK)* reaching its maximum at K_{NUC} = 5, suggested the existence of five (potentially six) distinct genetic clusters. Assuming K_{NUC} = 5, the first cluster grouped all *O. lehmanni* specimens, while the other four clusters corresponded to the mtDNA clades I-IV (Fig. 3-1d). There were 37 individuals exhibiting admixed ancestry (q = 0.1–0.9), and 73% of them were found in four populations (Angostura and El Salero n= 14 , Clade I *x* Clade II; La Choza and El Cocal, n= 13, Clade IV *x O. lehmanni*). The geographic location of these admixed populations (Fig. 3-1a) suggests that these localities represent two

different contact zones, one between the NW and the NE demes, and the second one between the South and the *O. lehmanni* ones. The results of the STRUCTURAMA analyses also supported the partition of K_{NUC}= 5 with a posterior probability of 74% (Table 3-1).

Table 3-1. Results from STRUCTURAMA analyses at a range of priors for shape and scale using the *O. histrionica* complex genetic dataset from microsatellite (n=6) and transcriptome-based markers (n=13). Optimal number of inferred populations is shown in light grey.

Priors used in STRUCTURAMA analyses						
Number of populations inferred	**Pr [K =i | X] Shape = 2.5 Scale = 0.5**	**Pr [K =i | X] Shape = 2.5 Scale =1**	**Pr [K =i | X] Shape = 5 Scale = 2**	**Pr [K =i | X] Shape = 10 Scale = 1**	**Pr [K =i | X] Shape = 10 Scale = 0.1**	**Pr [K =i | X] Shape = 1 Scale = 1**
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0.36	0.39	0.34	0.19	0.25	0.48
5	0.6	0.58	0.62	0.72	0.74	0.5
6	0.04	0.03	0.04	0.09	0	0.02
7	0	0	0	0	0	0

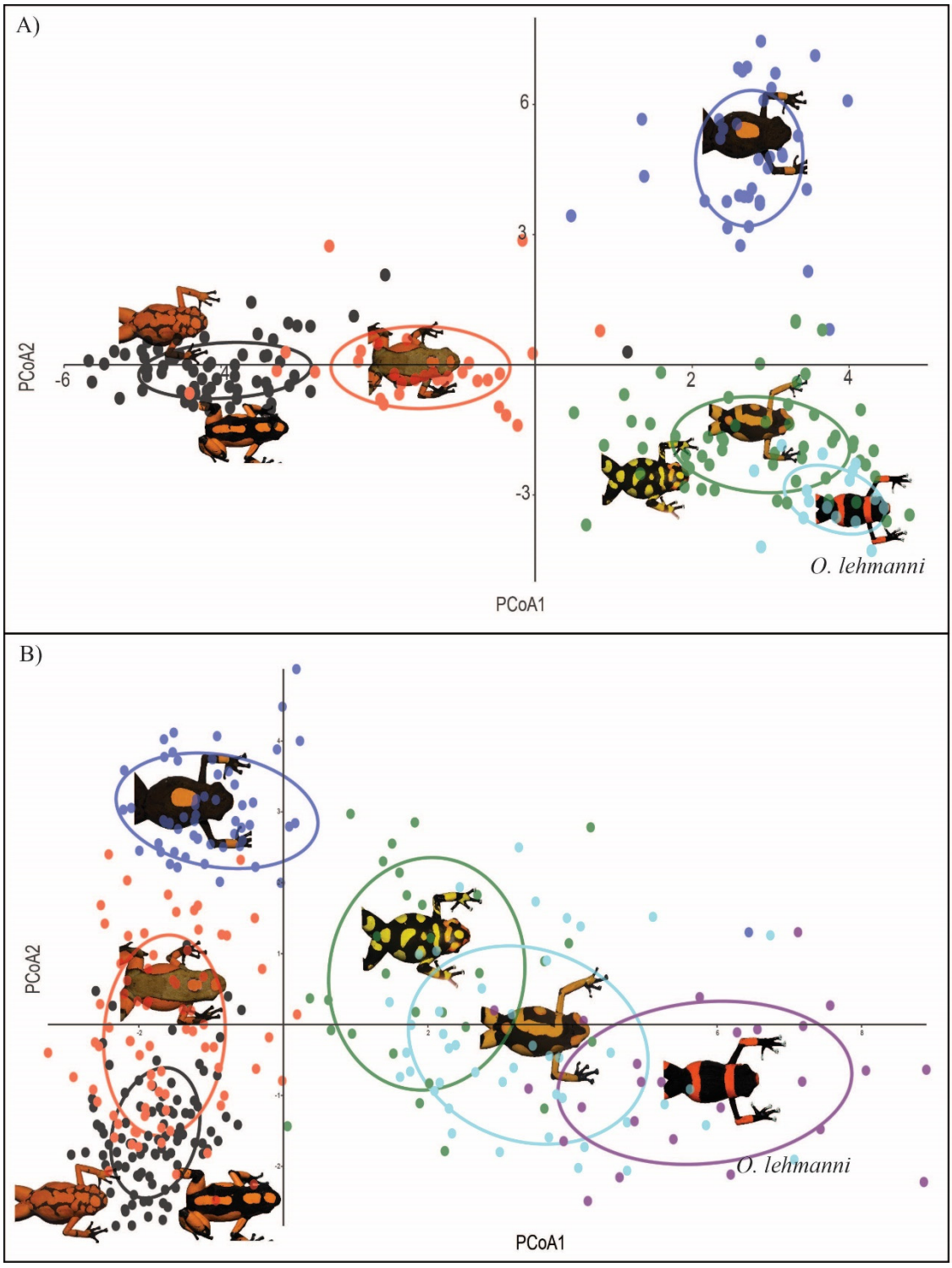


Figure 3-6. Genetic clusters by PCoA of allele frequency data from microsatellites (a) and transcriptome-based markers (b). Five cluster were detected by microsatellites markers while next generation sequencing data suggested the formation of six distinct clusters.

3.3.3 Lineage delimitation by integrative analysis

To first delimit the number of lineages, I used an integrated dataset composed of 12 multivariate dimensions (see Methods) and two different clustering methods: first, we used a model-based Gaussian clustering. The best model corresponded to an equal- covariance and diagonal distribution model (VEI) (Fraley & Raftery, 2002) with the BIC plateaued at K= 5 (Fig. S3C.1). Bootstrap likelihood ratio tests (*MclustBootstrap,* n= *999*) for K = 1-7 supported the existence of five lineages representing *O. lehmanni* (including the genetically admixed populations of La Choza – El Cocal), and the S, Central, NE, and NW lineages of "*histrionica*". Such taxonomic partitioning resulted in 100% assigned individuals (PP ≥ 0.95). The second-best model (K=6) split *O. lehmanni* into 'pure' (El Filo –La Cruceta) and admixed (La Choza – El Cocal) populations. Secondly, the hierarchical clustering approach was fully concordant with the Gaussian approach and showed five highly supported clusters (BS>95%) with identical composition of specimens (Fig. 3-7a). Given these results, I subdivided the *histrionica* complex into five provisional independent lineages. To test the robustness of these provisional delimitation, I performed a second round of clustering analyses in an independent set of specimens (n= 214) and K= 5. Under such delimitation, all individuals were correctly assigned to the expected lineage with PP ≥ 0.95 (Fig. 3-7b). Clustering analyses excluding mtDNA yield identical assignment results.

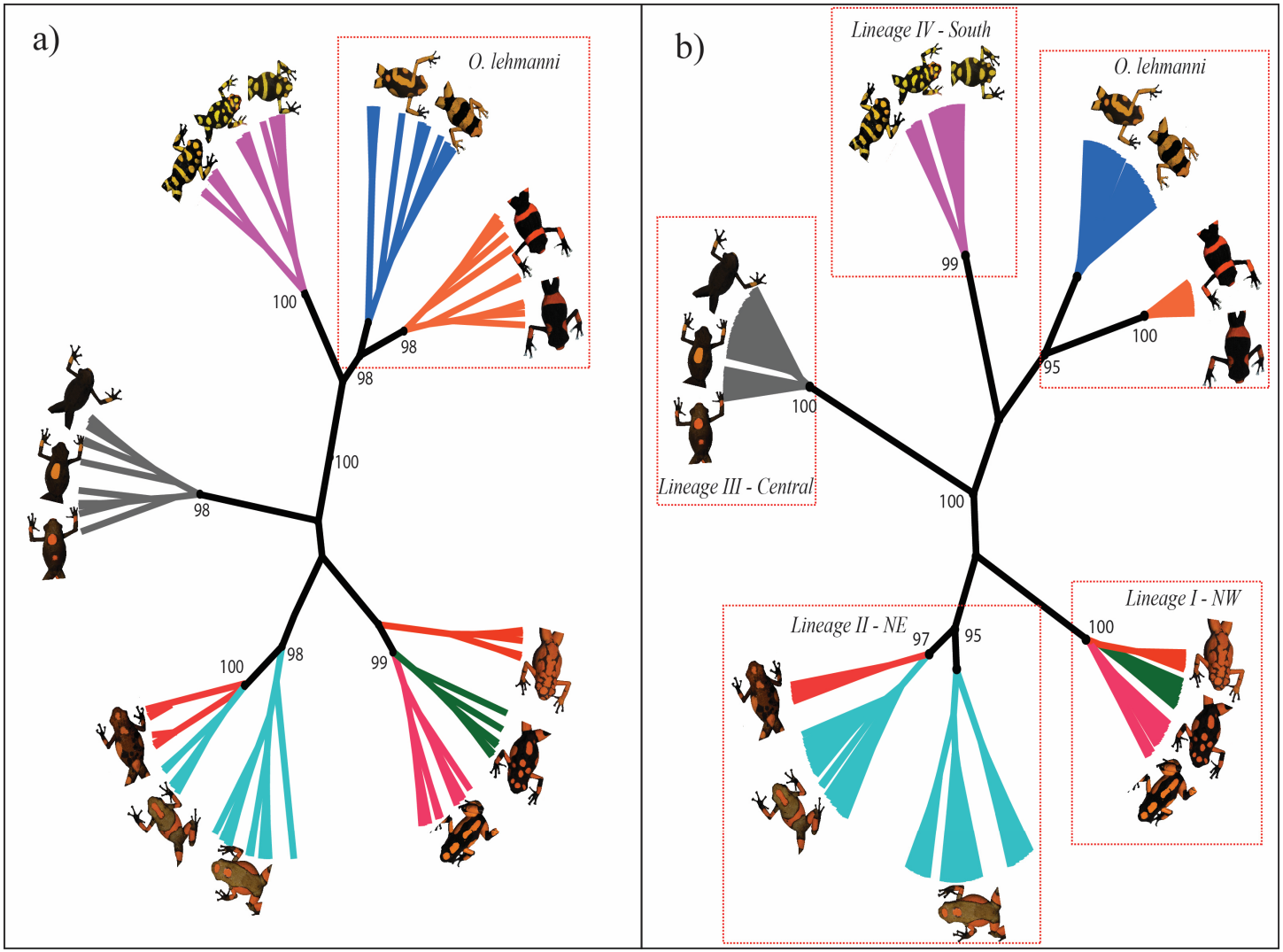


Figure 3-7. Integrative lineage delimitation based on the hierarchical clustering of standardized orthogonal variables (genetic, environmental and phenotypic). a) Provisional delimitation resulting from the analysis of 84 specimens. b) Extended clustering analysis of 214 individuals. The coloured boxes enclose individual specimens belonging to the same Gaussian cluster, PP ≥ 0.95). Only nodes with support >95% are indicated.

3.4 Discussion

3.4.1 Conceptual basis

Edwards and Knowles (2014) developed a statistical approach to evaluate hypotheses of species limits based on multivariate analyses and Gaussian clustering of total evidence (morphological, ecological and genetic) datasets. Here, I extended their approach and propose that the iterative implementation of such methodological framework can be used to delimit lineages under the GLC without any *a priori* species boundaries. The rationale here is that these multivariate clusters should correspond to evolutionarily independent metapopulation lineages because they reflect the common signal of different secondary defining properties (environmental and genetic distinctiveness, phenotypic diagnosability, etc.), implying the existence of barriers preventing or limiting gene exchange. It is important to recognize that because of the hierarchical nature of evolution, we expect clustering to occur at many levels. That is, we should observe a continuum of increasing degree of cluster overlap as the levels of genetic, phenotypic and environmental divergence decreases between species to variation across populations. Thus, depending on the focal system, the inferred clusters might represent inter-population variation, local host races, subspecies, species, etc. So, where along the divergence continuum separate lineages (clusters) should be recognized as distinct species?

To try to answer this question, one may rely on "arbitrary" statistical thresholds. For example, it has been argued that 70% correct multivariate assignment is a reliable threshold for species delimitation in complex groups (Valcárcel & Vargas, 2010). However, as noted by Coyne and Orr (2004, pg. 449) species do not have a greater "objective" reality than lower or higher taxa and we do not see a fundamental distinction between species and any other grouping

level. A better alternative is to contrast the multivariate clusters with those resulting from the application of multispecies coalescent methods that infer species status from a genealogical perspective.

In the case of the *histrionica* complex studied here, the GMYC analysis showed strong correspondence between the inferred multivariate clusters and the branching pattern of the mtDNA tree. The only source of discordance was that the integrative analysis grouped a few *O. lehmanni* specimens together with individuals from the '*histrionica'* mtDNA cluster IV (Fig. 3-4 and 3-5). Multivariate and Bayesian analyses showed that all 'misclassified' individuals are of admixed origin (Fig. 3-1b and 3-1d), a result that is concordant with previous studies suggesting hybridization between *O. histrionica* (*sensu lato*) and *O. lehmanni* (Medina *et al.*, 2013).

Despite its limitations, I consider that this approach has some practical advantages over other delimitation methods often reliant on genetic data. First, assuming that there is enough variation, this method provides a combination of intrinsic characters (attributes) that can be used to potentially recognize and describe newly discovered species. Second, because this approach relies in the cumulative effect of independent characters it is likely to be useful to unravel species in early differentiation stages were varying levels of differentiation across genetic, phenotypic and environmental axes are expected (Padial & De La Riva, 2010; Padial *et al.*, 2010). Third, this method can be useful for analysing large datasets. Most coalescent based delimitation methods could present computational challenges when analysing significant numbers of loci, putative species and/or individuals. In contrast, Gaussian clustering requires little computational power and can be easily scalable to include numerous genetic loci as well as large numbers of specimens.

3.4.2 Taxonomic implications

A key question in this study was to determine whether the extensive phenotypic diversity within the *O. histrionica* complex reflects the existence of several independent taxonomic units *(i.e*. species) or just one (possibly two) highly polytypic species (Myers & Daly, 1976; Grant *et al.*, 2006). The application of the integrative multivariate procedure supported the taxonomic status of *O. lehmanni* and provided statistical support for the existence of five distinct independent evolutionary lineages of *Oophaga* which deserve further research and with potential to be formally recognized and described as new species.

Though the specific status of *O. lehmanni* has casted doubt among taxonomist during the last decades (Lötters, 1992; Lötters *et al.*, 1999; Medina *et al.*, 2013), the present analyses indicated a considerable amount of genetic, environmental and phenotypic differentiation between the two current nominal species *O. lehmanni* and *O. histrionica* (Myers and Daly, 1976; Fig. 3-1e). Thus, this study supported the taxonomic status of *O. lehmanni*. This study also provides further support for the existence of a geographically isolated hybrid lineage (i.e. cluster) of *O. lehmanni* (Medina *et al.*, 2013). Overall, the pattern of genetic additivity of parental marker alleles (Fig. 3-1d), along with environmental (Fig. 3-1c) and phenotypic separation (Fig. 3-1b) suggests that hybridization between these two lineages may have resulted in the successful establishment of a distinct lineage. Further genetic studies are required to test this hypothesis.

Here, I also propose the existence of four independent evolutionary lineages within *O. histrionica* (*sensu lato*). These lineages correspond to each of the four '*histrionica'* clusters revealed by the integrative analyses. Although detailed morphological descriptions have been previously proposed (Berthold, 1846; Silverstone, 1975; Myers & Daly, 1976; Daly *et al.*,

1978; Lötters, 1992; Lötters *et al.*, 1999; Grant *et al.*, 2006), I have included a summarized description of each lineage with respect to body measurements and a synthesized description of colouration patterns based on the phenotypic analysis. I would like to highlight that if we accept "species" as independent evolutionary units, the new delimited *Oophaga* lineages (Fig. 3-7b) have the potential to be further and formally recognized as new species to science. Several lines of evidence and different methodological approaches supported the lineages described in this Chapter. However, the inclusion of other lines of evidence (e.g. acoustic, behavioural, etc.) over an extended geography may result in new, revised delimitations.

Oophaga histrionica **[Lineage I - NW**]. This lineage is supported by the mtDNA (PS=1), nDNA and the general conservative agreement of the integrative analysis (BS=99%). Museum specimens of this lineage are deposited at the American Museum of Natural History (AMNH) under accession numbers 86979-86985 (Myers & Daly, 1976).

Range: Based on our own sampling plus data collected from different authors (Silverstone, 1975; Myers & Daly, 1976; Mendez-Narvaez & Amezquita, 2014), populations of this lineage inhabit the coastal lowlands at the west banks of the San Juan and Atrato rivers which, due their width, presumably represent a major barrier to gene flow between the eastern and western *Oophaga* lineages (Fig. 3-1a). Populations are distributed in a relatively wide geographic extension of dense rainforest with known populations from Mecana, Bahia Solano (Lat. 6.29640, Lon.-77.37869) and Serranía del Baudo (Silverstone, 1973) throughout La Victoria, (Lat. 5.51074, Lon. -76.86979), El Salero, (Lat. 5.36412, Lon.-76.64629), Nuqui, (Lat.5.70652, Lon.-77.26079) until Quebrada Docordo, department of Chocó (Lat. 4.56265, Lon.-77.01399). In contrast to *O. histrionica* [Lineage III – Central), this lineage inhabits lower altitudes (0 to 324 m.a.s.l) with an annual mean temperature of 25.9 ºC and annual mean

precipitation of 6279 mm.

Phenotypic description: Body size in these frogs ranges between 53-33 mm (36.2 ±1.7 mm, n=75). This lineage represents three different phenotypic clusters (morphotypes) detected in this study with high bootstrap support (BS>82%) (Fig. 3-2). Individual frogs are black with bright orange or reddish orange spots that are highly variable in size and number (n=3-30); however, the coloured dorsal area remains relatively constant (40-70% of total dorsal area). Colour variation is exemplified by Myers and Daly (1976) (Fig. 10, pp. 209).

Oophaga histrionica **[Lineage II - NE**]. This lineage is supported by the genetic delimitation (mtDNA and nDNA, PS=1) and the general conservative agreement of the integrative analysis (BS=98%). Museum specimens representing this lineage are AMNH 85186-85189 and AMNH 86957-86960 (Myers & Daly, 1976)

Range: This lineage is distributed at the east banks of the Atrato river which probably represent a geographic barrier between *Oophaga* Lineages I and II (Fig. 3-1a). After confirmation of the presence of this lineage in those geographic locations originally visited by Myers and Daly (1976) plus our own field records, I consider that populations of this lineage are distributed from the surrounding areas of Tutunendo, Chocó (Lat.5.75219, Lon.-76.53634) throughout Pacurita, Chocó (Lat 5.646565, Lon.-76.56655), Bagado, Chocó (Lat.5.38121, Lon.-76.42759) until Vereda La Angostura, Playa de Oro, Chocó (Lat. 5.32194, Lon.-76.43090). Similar to Lineage I, individuals inhabits the relatively warmer lowlands (annual mean temperature of 26.6 ºC; altitude 55 to 168 m.a.s.l). However, Lineage II exhibits the higher requirements of humidity, inhabiting locations with annual mean precipitation >7500 mm.

Phenotypic description: Frogs of this lineage are similar in body size to *O. histrionica*

lineage I (34.9 ±2.4 mm, n=43) and are represented by two distinct morphotypes detected in this study (BS>91%) (Fig. 3-2). The basic and most common colouration pattern is represented by one to three large orange-reddish-yellow spots and bracelets on a background ranging from light to dark brown, but never black in colour. Detailed description of variation in spot sizes and distribution, as well as a representation of morphologic variation is presented in Myers and Daly (1976) as "*Playa de Oro*" frogs (Fig. 8-9, pp.206-208).

Oophaga histrionica **[Lineage III - Central]**. This lineage was originally described as a species (*Dendrobates histrionicus*) (Berthold, 1846) based on descriptions of an unknown holotype (reproduced in Fig. 6, pp. 200, Myers & Daly, 1976) phenotypically similar to frogs occurring on the upper Rio San Juan, in the Department of Risaralda, Colombia. In this study, this lineage was highly supported by all individual lines of evidence and the integrative analysis (Fig. 3-1e). A detailed description of colouration and pattern variation of this lineage is presented in Myers & Daly (1976) as "*Santa Cecilia*" frogs (pp. 205). Museum specimens that represent this lineage are deposited at the AMNH under collection numbers AMNH 85159-85161 (Myers & Daly, 1976).

Range: Based on previous reports an my extensive field work, this lineage is distributed in a narrow geographic range at an altitude between 342-410 m (mean 394 m). Populations are geographically restricted to less than 30 km^2 in the surrounding areas of the municipality of Santa Cecilia, Colombia (Lat. 5.30336, Lon. -76.21662) with annual mean temperature of 25.6 ºC and annual mean precipitation of 5267 mm.

Phenotypic description: Frogs of this lineage are significantly smaller in body size than other lineages within the complex (32.9 ±1.7 mm, n=28). They represent a clearly defined and unique phenotypic cluster (BS=82%) (Fig. 3-2) where individual frogs have light brown to

black bodies with usually one or two large orange or red-orange spots on the back. In rare cases, the dorsal marking is absent or difficult to notice. Morphological variation is shown in Myers and Daly (1976) (Fig. 7, pp. 205).

Oophaga histrionica **[Lineage IV- South**]. This lineage was supported by the genetic delimitation (mtDNA and nDNA, PS=1) and the general conservative agreement of the integrative analysis (BS=100%). Museum specimens representing this lineage are deposited at the Universidad del Valle, Cali (UVC 12431-52) (Lötters *et al.*, 1999), at the Institut Royal des Sciences Naturelles de Belgique (IRSNB) (MRHN 1038) and AMNH 10613-15 (Silverstone, 1975).

Range: This lineage is located at southern localities (Fig. 3-1a) and its geographic range correspond to <170 km^2 at Valle del Cauca department, Colombia, at higher elevations when compared with other *Oophaga* lineages (616-791 m.a.s.l). Populations have been reported at Delfina, Buenaventura (Lat.3.81364, Lon.-76.81305), Cisneros (Lat.3.79955, Lon.-76.80305), Rio Zabaletas, Buenaventura (Lat.3.72003, Lon.-76.91802), Bajo Anchicaya, Dagua (Lat.3.63621, Lon.-76.94299), El Danubio, Dagua (Lat.3.62210, Lon.-76.90199) and vereda El Cauchal, Dagua (Lat.3.61712, Lon.-76.865). In contrast to lineages I and II, this lineage inhabits cooler and dryer geographic regions (annual mean temperature=24.4 ºC; annual mean precipitation=2419 mm).

Phenotypic description: These frogs are significantly bigger in body size than other *O. histrionica* lineages (41.2 ±1.3 mm, n=29). Background colouration is always completely black with dorsal yellow, red-orange or orange spots highly variable in number (3 to 20) and size. Dorsal pattern sometimes includes median bracelets as *O. lehmanni* (Myers & Daly, 1976) that are incompletely formed ventrally in most cases. Some populations as "*Delfina*" and

"*Cisneros*" (Silverstone, 1975; Medina *et al.*, 2013) present an increased reddish colouration towards the head and mouth (first frontal quarter of size length). This pattern is locally known as "*red head*" frogs and is shown in Lötters *et al.* (1999) (Fig. 3, pp. 27). However, this lineage is represented by a single morphological cluster with high statistical support (BS=95%) (Fig. 3-2).

3.4.3 Conservation implications

This study also illustrates how understanding lineage diversification can inform wildlife management and has relevance to the conservation efforts of South American poison frogs. The integrative lineage delimitation proposed here has very important conservation implications as it revealed that some of the *O. histrionica* lineages should be considered amongst the most vulnerable of the Neotropical frogs: of the five lineages (*O. lehmanni* and *O. histrionica* lineages I-IV) two of them (*O. lehmanni and O. histrionica* lineage III) are single-locality endemics and should be both considered critically endangered. The inclusion of *O. lehmanni* as a critically endangered species more than 10 years ago led to the proposal of several conservation programs and regional legal policies aiming to preserve the remaining populations and their habitat (Bolivar *et al.*, 2004; Valencia-Zuleta *et al.*, 2014). This study revealed that the distribution of *O. histrionica* lineage III is restricted to a very small area (< 30 km^2), and the only three known populations of this microendemic lineage should probably be subject of similar actions. Also, given that individuals of these two lineages are highly sought-after in the pet trade market, it will be desirable to include them in the Appendix I of the CITES treaty to prevent the commercial trade of wild-caught specimens.

The conservation status of the other *O. histrionica* lineages (I, II and IV) is less clear as they all have relatively wide, discontinuous, distributions. Although their effective

population sizes are likely to be large, preliminary data suggest that populations are strongly structured. Therefore, several management (MUs) and evolutionary significant (ESUs) units (Funk & Richardson, 2002) should probably be considered. In the absence of any better information regarding species population ecology, density or abundance, we consider that these lineages should be temporarily categorized as vulnerable to extinction. Regardless of their conservation status, none of the lineages described in this paper occur in protected areas. Ideally, once included in the appropriate endangered list, these charismatic frogs will serve as umbrella species for the future conservation of entire ecosystems.

CHAPTER 4

COMPARATIVE TRANSCRIPTOMICS OF HARLEQUIN POISON FROGS REVEALS COLOURATION AND TOXICITY – RELATED GENES

4.1 Introduction

A wide range of organisms use aposematic, or warning, colouration to signal their unprofitability to potential predators (Harvey *et al.*, 1982; Brown, 2013; Cummings & Crothers, 2013). Amongst them, the poison frogs (Dendrobatidae) from the tropical rain forests of Central and South America possibly represent the most spectacular example. While the majority of Dendrobatids rely on crypsis to avoid predators, some members of this group are brightly coloured and chemically defended (Grant *et al.*, 2006). Comparative analyses have revealed that aposematism in Dendrobatids has likely evolved more than once resulting in clades with aposematic, cryptic, and/or mimic species (Santos *et al.*, 2003): while in some clades the colouration of harmless frog species has evolved to mimic that of the toxic ones (Batesian mimicry) (Mallet & Joron, 1999; Wüster *et al.*, 2004; Holen, 2013) in some other groups the colouration of chemically defended species has coevolved to produce a similar warning signals (Mullerian mimicry) (Sherratt, 2006; Sherratt, 2008). Hence, Dendrobatids represent an ideal system to dissect the genetic bases of aposematic related traits and —more generally speaking— to explore the molecular evolutionary processes underlying phenotypic convergence and diversification.

From a genetic point of view, aposematism can be defined as a complex phenotype resulting from the integration (*i.e*. covariation) of different genetic elements related to conspicuousness, bold behaviour, unpalatability, diet specialization, etc. While a long history of research has been devoted to understand the genetics of warning colouration in arthropods,

particularly in *Heliconious* butterflies (Ramos & Freitas, 1999; Langham & Benkman, 2004), the molecular underpinnings of aposematism in vertebrates, particularly the mechanisms whereby individuals become toxic (or distasteful) remain mostly unknown. In this study, my primary aim was to shed light on the molecular genetic bases of three main components of aposematism in poison frogs: the ability to sequester alkaloid-based chemical defences, warning colouration and the auto resistance to avoid self-intoxication. To achieve this goal, I performed a comparative analysis of the skin transcriptome of Harlequin poison frogs (*Oophaga histrionica* species complex). This study used data from four *O. histrionica* lineages (I – IV*)* (Posso-Terranova & Andrés, 2016b) plus the currently recognized *O. lehmanni;* all of them inhabiting the lowland Pacific rainforests of the Colombian and Ecuadorian Chocó. Preliminary molecular data from nuclear and mitochondrial markers showed a highly similar genetic background among lineages in contrast with an extraordinary diversity of morphotypes (Medina *et al.*, 2013). Individuals from different populations can either relatively homogeneous, striped, or spotted, and their colours range from bright red, to orange and yellow. These conspicuous colouration patterns serve as a warning signal of their chemical defences, a complex mix of diet-derived alkaloids secreted by the dermal glands (Saporito *et al.*, 2006; Saporito *et al.*, 2012).

Here, I hypothesized that the genes, pathways, and/ or gene networks potentially associated with colouration, alkaloid metabolism, transport and storage, should be highly expressed in skin tissue of all the studied species. Because these lineages are closely related, I did not expected a high turnover rate (gain and/or loss of orthologs) across the genus. The alkaloids of these frogs are known to disrupt the normal ion-channel activity and to alter the neurotransmitter-receptor binding capacity of nerve and muscle cells (Daly *et al.*, 1978; Daly *et al.*, 2002; Daly *et al.*, 2003). Therefore, I expected these species to contain adaptations to deal

with auto toxicity. Metabolic detoxification is one the main mechanism underlying alkaloid resistance (Santos *et al.*, 2016). Thus, I also hypothesized that conserved detoxification genes (e.g. cytochrome P450), should be highly expressed in the skin of these frogs.

Aposematic organisms frequently show strong luminance contrast among different elements of their colouration (e.g. the black and yellow striping of wasps, the black and white patterning of skunks, etc.). Such mottled patterns are thought to increase aversion learning of an aposematic signal relative to solid bright colours alone (Qvarnström *et al.*, 2014). Although the *Oophaga* species studied here display a wide variety of warning signals, all but one (*O. sylvatica*) share a black background colouration (Fig. 4-1). In anurans, colouration relates to both the structure and pigment composition of the dermal chromatophore units in which three cell types (xantophores, iridiophores, and melanophores) are laid one on another (Bagnara *et al.*, 1968; Bagnara & Hadley, 1969; Bagnara *et al.*, 1979). In dark skin areas, the distal fingers of the melanophores are filled with melanin granules (melanosomes) obscuring the pigment of the other cells of the unit. Thus, I also hypothesized that genes involved in the amount, size and distribution of the melanosomes are highly expressed in the skin of these frogs.

To test these hypotheses, I investigated the common transcriptional profiles of several *Oophaga* lineages (See Chapter 3). For each of the lineages, more than 18,000 ESTs were generated through Illumina sequencing (Roche) and assembled in more than 30,000 contigs. To our knowledge, these represent the first transcriptome dataset for Dendrobatid frogs. Functional annotation and comparative analyses were performed to identify candidate genes. The identification of more than 250 orthologous contigs across species allowed me to test the current phylogenetics hypothesis for this group and to screen for differential substitution rates along lineages and for individual genes. Overall, this study provides an important molecular resource

for the study of aposematism within Dendrobatids and will facilitate the assembling and annotation of future Dendrobatid genomes.

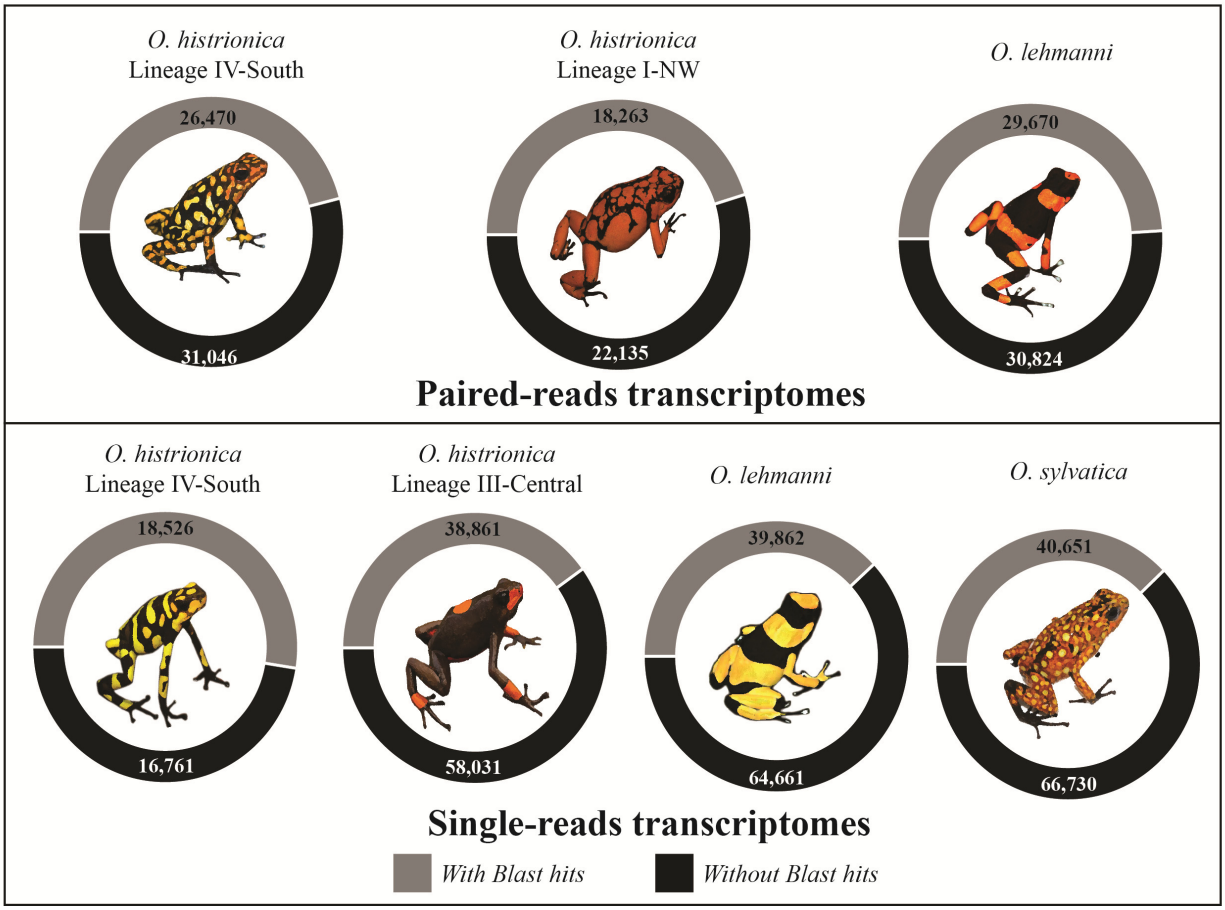


Figure 4-1. Pie charts representing the number and proportion of contigs with significant *Blast* hits ($E<1.0E^{-5}$) for each individual transcriptome (n=7). *O. histrionica* lineages correspond to those described in Chapter 3 and indicated in Fig. 3-1.

4.2 Materials and Methods

4.2.1 Library preparation and sequencing

Seven individuals (*O. lehmanni, n=2; O. histrionica, n=4;* and *O. sylvatica,* (Funkhouser, 1956; Myers & Daly, 1976) were euthanized in the field with benzocaine gel at 5%. Individual skin samples were taken and stored at -80 °C in RNAlater® (Life Technologies,

Carlbad, CA) as soon as possible. Individual RNA extractions were performed using Tri-Reagent (Molecular Research Center, Cincinnati, OH). After quality check (Agilent Bioanalyzer Agilent Technologies, Wilmington, DE) samples were used to generate RNAseq libraries using the Illumina Truseq RNA Sample Prep protocol (Illumina, San Diego, CA). Libraries were cleaned using AMPure XP and sequenced on a single Illumina HiSeq2000 lane (TruSeq SBS v. 3) as follows: four libraries (*O. histrionica=2, O. lehmanni=1*, and *O. sylvatica*=1) were run in a lane of single-end (150-mer) while the other three (*O. histrionica=2* and *O. lehmanni=1*) were run in a single lane of a paired-end module (100-mer, x2). All animal procedures were approved by the Ethic Committee of the Universidad Nacional de Colombia (Acta No.03,) and were conducted based on the NIH Guide for the Principles of Animal Care. Sequencing was carried out by the Cornell University's BioResource Center.

4.2.2 Transcriptome assemblies and functional annotation

Initial read quality trimming, filtering and removal of adapters was carried out using FLEXBAR v.2.5 (Dodt *et al.*, 2012). All retained reads were ≥ 50bp with an average quality of ≥ 30, and less than two uncalled (ambiguous) bases (Table S4A.1). For each library, a *de novo* transcriptome was assembled using TRINITY v.2.1.0 (Haas *et al.*, 2013) with default settings (kmer=25, minimum contig length=48) keeping only the longest transcript per cluster for subsequent analyses. To construct a composite *de novo* reference transcriptome, I followed Gould *et al.* (2015) and combined 20% randomly selected reads (>33 million in total) from each of the four single-end libraries (*O. histrionica=2, O. lehmanni*=1, *O. sylvatica*=1). Then, a TRINITY assembly was performed using a minimum contig length of 350 bp. To filter out highly similar contigs that may represent alternatively spliced transcripts, I implemented the error correction module of iAssembler v1.3.2 (Zheng *et al.*, 2011) with default parameters

(maximum length of end clips=30 bp, minimum overlap length=40 bp, minimum percent sequence identity=95%).

Gene annotation was conducted using sequential *blastx* search to both the available *Xenopus* transcriptomes and the NCBI nr database. Briefly, the composite transcripts were first compared with that of the *Xenopus* databases retaining annotations with E-values ≤ $10^{-5.}$ Unannotated contigs were then submitted for *blastx* to the non-redundant (nr) protein database of the NCBI for possible identification. gene ontology (GO) annotation and term mapping was done by using Blast2Go using default significance cutoffs (Conesa *et al.*, 2005).

To estimate the completeness of each transcriptome, I implemented the Core Eukaryotic Genes Mapping Approach (CEGMA) (Parra *et al.*, 2007; Parra *et al.*, 2009) as implemented in TRUFA (Kornobis *et al.*, 2015). This analysis assesses the quality of any given transcriptome by estimating the completeness of 458 core genes predicted to be ubiquitous in eukaryotes. The pre-processing of the reads and all *blastx* analyses were run in the Bugaboo Dell Xeon cluster of the western Canada's WestGrid computer facilities (www.westgrid.ca).

4.2.3 Highly expressed unigenes

I performed further analysis of read distribution and the relative abundance of different transcripts in each of the source tissues by mapping RNA-sequencing reads back to the composite *de novo* assembly. In order to identify highly expressed unigenes in *Oophaga* skin tissue, I first implemented eXpress (probabilistic assignment of ambiguously mapping sequenced fragments) (Roberts & Pachter, 2013) to estimate the effective number of reads (ER) that mapped to the contigs in the reference transcriptome after adjusting for read number and length biases. Then, to compare the proportion of reads that mapped to a transcript in each RNA-library (n=7), I estimated the number of Transcripts per Million (TPM), a measure of

RNA abundance that allows the comparison between samples (sum of all TPMs in each sample are the same) (Wagner et al., 2012). I visually explored for highly represented unigenes in each RNA-library by using dispersion plots of the ER values and percentile plots of the TPM distribution. I performed a Principal Component Analysis (PCA) of the reference contigs dataset and using each library TPMs values as independent variables (n=7). I then selected the top expressed unigenes (2%) for further *blast* analysis and GO term mapping.

4.2.4 Clustering of orthologous sequences

To obtain sequences suitable for a comparative analyses, I generated two datasets which included orthologous ESTs across the studied lineages. I implemented a workflow similar to that of Chamala *et al.* (2015) that allowed the identification of conserved single copy nuclear SNC orthologous present in all individual libraries. To do so, I used the CEGMA/reciprocal Blast-based pipeline described in Figure 4-2. Then, I used reciprocal *Blast* (*tBlastn – Blastx*) between the set of identified CEGs (n=362) and each of the individual transcriptomes to identify common SCN orthologous. Filtering and clustering of CEGs orthologous present in all seven transcriptomes were performed by using a customized *Perl* script (available upon request). The final steps included the re-orientation of reversed contigs by using the offline version of the Sequence Manipulation Suite (Stothard, 2000). This approach resulted in the identification of 362 orthologous groups (dataset #1). Because this approach is limited to highly conserved genes, I generated a second dataset using OrthoMCL (Li *et al.*, 2003) and default parameters. Groups missing one or more lineages were discarded. If there were multiple transcripts from the same lineage within a group, the one with the highest sum of BLAST bit scores to other lineage's transcripts was selected. This strategy resulted in the identification of 4,237 (1:1:1:1:1:1) groups of putative orthologs (dataset #2).

4.2.5 Phylogeny and genetic distances

For the dataset #1, the resulting supermatrix alignment was used to estimate the best model of nucleotide substitution using the sample size-corrected Akaike Information Criterion (AICc) (Posada & Buckley, 2004) as implemented in jModelTest v2.1.4 (Posada, 2008). DNA sequence variation estimates were calculated using DNAsp v.5.10.01 (Librado & Rozas, 2009) and genetic distances were estimated using the best model of nucleotide substitution in MEGA6 (Tamura *et al.*, 2013). Phylogenetic trees were constructed using maximum likelihood (ML) and Bayesian inference (BI) as follows: first, using MEGA6 and the results derived from jModelTest, I obtained an ML tree using extensive (level 5) Subtree Pruning and Regrafting (SPR) heuristic searches. The relative support for each node was estimated by generating 1,000 bootstrap replicates. Second, I constructed a BI-based tree using BEAST v. 2.1.2 (Drummond *et al.*, 2012; Bouckaert *et al.*, 2014) with a relaxed molecular clock and an uncorrelated log-normal (UCLN) model of molecular rate heterogeneity. I ran three chains for 10 million generations sampled every 1000 steps. The resulting trees and log files of the three independent runs were combined using LogCombiner v. 2.1.2. (Bouckaert *et al.*, 2014). For each estimated parameter, convergence was assessed using Tracer v. 1.6 (http://tree.bio.ed.ac.uk/software/tracer/) and effective samples sizes (ESS) were calculated to ensure adequate mixing (ESS>300, after 30% burn in). I summarized the posterior probability density of the combined tree and log files as a maximum clade credibility (MCC) tree using TreeAnnotator v. 2.1.2 (Bouckaert *et al.*, 2014). For dataset #2, I concatenated each transcriptome CEGs dataset to "supergenes" and performed a multiple sequence alignment of the seven supergenes in MAFFT (Katoh *et al.*, 2002; Katoh & Standley, 2014) to obtain a final super-matrix for further phylogenetics analysis. As a secondary estimate of topology and to

better determine branch lengths, I also undertook a tree estimation using the larger dataset (#2) and a concatenation-based ML approach in RAxML8 using the GTRCAT model and 1,000 bootstrap replicates. All analyses were run in the Breezy Dell Xeon cluster of the western Canada's WestGrid computer facilities (www.westgrid.ca).

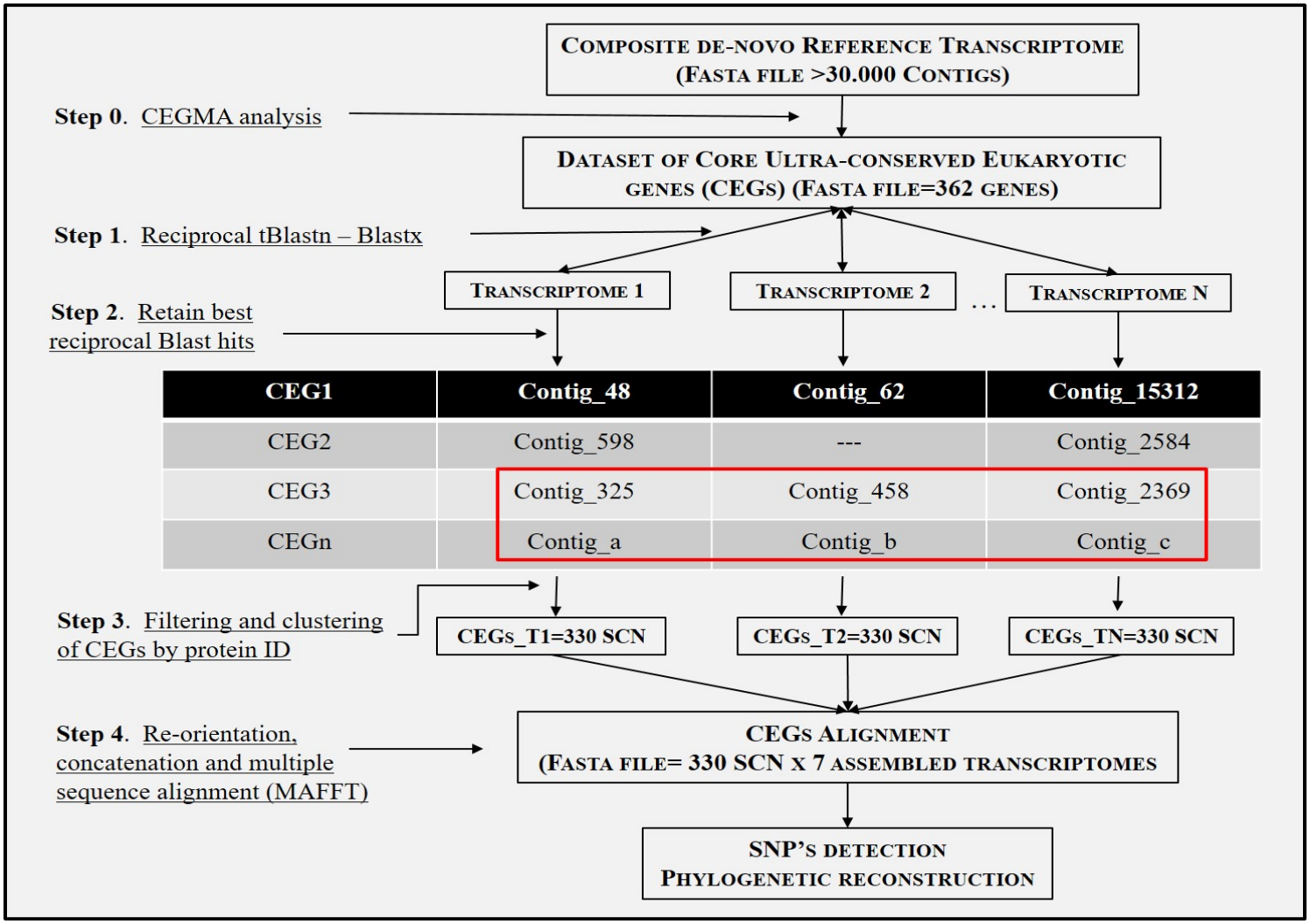


Figure 4-2. Flow-chart indicating the different steps involved in the approach for single copy orthologous (SCN) discovery and further analysis. Detailed information about bioinformatics tools required on each step is provide in the methods section.

4.3 Results

4.3.1 Transcriptome assemblies and functional annotation

Illumina sequencing produced an average number of reads per sample of 118.3 million for paired-end and 37.7 million for single-end libraries. After a fairly stringent read trimming involving the removal of low quality sequences, duplicated reads and reads containing adapter/primer sequences, I retained over 81% of the initial sequence data in all cases. Paired-end libraries produced a fairly consistent number of reads while only one library (*O. histrionica*-south) produced < 10 million reads (Table S4A.1). Datasets including raw RNA sequence reads for all seven individuals will be deposited to the Genebank Short Read Archive prior to manuscript publication.

After the removal/merging of highly similar sequences representing spliced transcripts (paired-end:7,5% - 9,6%; single-end: 6.0% - 13.4%), our individual transcriptomes recovered a large amount of contigs ranging from 35,287 (*O. histrionica*-south) to 107,381 (*O. sylvatica*). Paired-reads libraries (100-mer; 2X) produced transcriptomes with number of contigs from 40,398 (*O. histrionica*-NW) to 60,494 (*O. lehmanni*) while in general, the longer single-reads (150-mer) produced transcriptomes with a higher number of contigs (>95,000) with the exception of the *O. histrionica*-south library (35,287 contigs) (Table 4-1). N50 values and average transcript length (AL) were lower for paired-end libraries (N50= 498-561bp; AL= 441.78±379 - 475.93±436bp) than for assemblies produced with single-end libraries (N50=667-1579bp; AL=538.9±1001bp - 668.6±819 bp) (Table 4-1). Comparative *blast* analysis indicated that our transcriptomes recovered a significant proportion of the *X. laevis* and *X. tropicalis* transcriptomes (50,592 and 41,042 unigenes respectively). Paired-end transcriptomes showed a lower number of significant *blast* hits (18,263 – 29,670) than single-

end libraries (18,526 – 40,651) (Fig. 4-1). In general, I obtained individual *de-novo* transcriptomes of good quality that recovered 37-89% of the complete and 60-92% of the partial protein count of the core eukaryotic genes (CEGMA) (Parra *et al.*, 2007; Parra *et al.*, 2009) (Table S4A.2).

The final *de novo* composite skin transcriptome yielded 31,498 contigs greater than 350 bp reconstructed from across skin RNA libraries of four individuals (*O. histrionica* South and Central, *O. lehmanni* and *O. sylvatica*). The overall assembly incorporated 94% of all initial reads and the level of fragmentation was low with half of all base pairs clustered into contigs of 1316 bp in length or greater. The maximum contig length was 16,041 bp and the AL was 1014.91 ± 948.43 (Fig. 4-3a). Nucleotide-based *blast* analyses (*blastx*) revealed that ~63% of the contigs (n=19,733) show significant similarities with either annotated gene products and/or known protein domains (E-value ≤ 10−5) (Fig. 4-3b) and only a small fraction of unigenes (3.4%) showed significant homology to the same annotated transcript. The maximum number of homologies was found with *X. tropicalis* (74%), *X. laevis* (2%) and the green sea turtle *Chelonia mydas* (0.5%). The composite skin transcriptome was of higher quality when compared with individual assemblies, having 83.3% of the complete and 92.7% of the partial protein count of CEGMA, with 53.5 – 61.95% having orthologous (Table S4A.2)

Table 4-1. Summary statistics and quality assessment estimators of seven individual and the composite reference transcriptomes generated in this study.

	Paired-end libraries			**Single-end libraries**				
	***O. histrionica* (NW)**	***O. histrionica* (South)**	***O. lehmanni***	***O. sylvatica***	***O. lehmanni***	***O. histrionica* (South)**	***O. histrionica* (Central)**	**Reference transcriptome**
File size (Mbytes)	22.23	32.06	35.7	93.9	77.3	23.16	85.97	31.6
# contigs	43,690	63,520	66,937	123,916	117,082	37,552	111,823	31,499
# contigs after filtration (removal/merging of highly similar contigs)	40,398	57,516	60,494	107,381	104,523	35,287	96,892	31,499
# bases	19,853,219	28,061,593	31,856,993	81,266,530	66,367,430	20,238,091	74,768,159	31,968,628
# bases after filtration	17,164,733	23,756,585	26,287,312	58,224,462	50,845,151	15,278,290	53,582,406	31,968,628
Merged contigs	7.5%	9.5%	9.6%	13.3%	10.7%	6.0%	13.4%	0.0%
Average transcript length (AL)	454.41± 405	441.78± 379	475.93 ± 436	655.8 ± 799	566.8± 648	538.9 ± 1579	668.6 ± 819	1,014.91 ± 948.43
Maximum length	6,115	7,915	8,400	15,102	12,906	54,265	15,136	58,024
N50	528	498	561	1079	809	667	1,101	1,316
Contigs with *Blast* hits	46.5%	47.7%	50.8%	42.1%	41.6%	53.4%	44.5%	62.6%

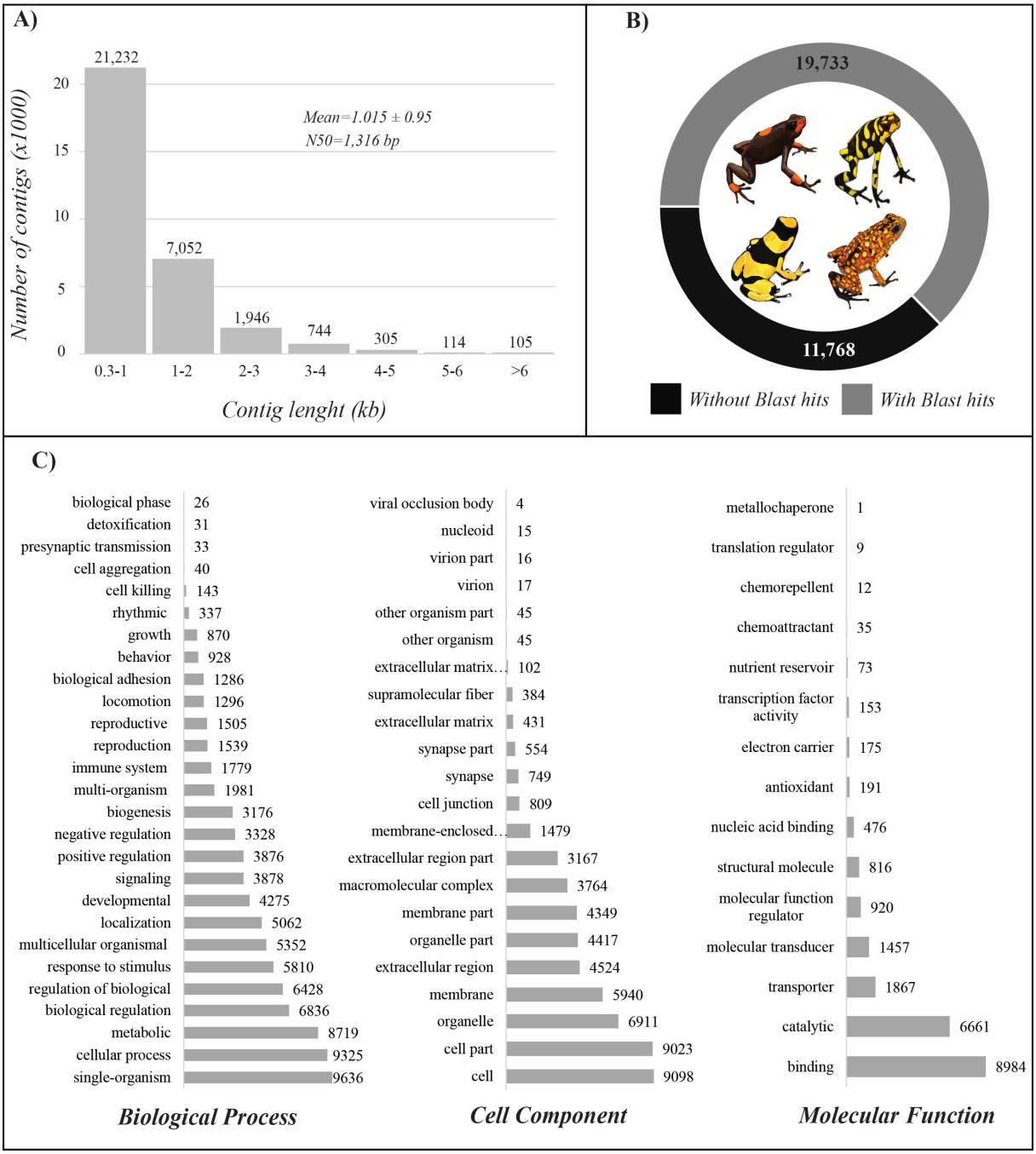


Figure 4-3. a) Contig length distribution of the *de-novo* composite reference transcriptome of *Oophaga* species from skin tissue. b) Pie chart representing the number and proportion of contigs with significant *Blast* hits ($E<1.0E^{-5}$) in the reference *Oophaga* transcriptome. c) Gene ontology (GO) categories distribution (level II) for the annotated unigenes in the *Oophaga* reference transcriptome.

Gene Ontology (GO) assignments were used to classify the functions of the predicted unigenes based on contigs with significant *blastx* (E-value ≤ 10−5). Based on GO level II,

unigenes were assigned to 27 biological processes (BP), 22 cell components (CC) and 15 molecular functions (MF) (Fig. 4-3c). Some unigenes were associated with multiple GO annotations because a single sequence may be annotated in any or all categories, giving more GO annotations than sequences annotated (Xie *et al.*, 2002). Within the BP, ~47% of the annotations were assigned to basic cellular, metabolic processes and biological regulation. Remaining unigenes were involved in a broad range of BP such as response to stimulus (7%), response to stress (7%), localization (6%), developmental process (5%), signal transduction (4%), biogenesis (4%), immune response (2%), reproductive process (1.7%) and cellular adhesion (1%). Within the CC category, other than the essential cell constituents (44%), the membrane structure components were highly represented in the transcriptome (19%). Within the MF, most of the unigenes were assigned to binding and catalytic activities (72%) followed by transporter activity (9%), molecular transducer (7%) and molecular function regulators (4%) (Fig. 4-3c).

4.3.2 Highly expressed unigenes

After adjusting for library size and length bias, the maximum effective number of reads (ER) varied between RNA libraries from 2,460.74 (*O. histrionica* south – single-end) to 5,812.90 (*O. sylvatica* – single-end) and corresponded to a highly expressed 28S ribosomal RNA gene. Dispersion plots of ER values showed a common pattern of over-represented unigenes across RNA-libraries. That is, the same group of unigenes showed similar patterns of over-representation no matter RNA library's origin (i.e. different individuals) or composition (paired vs single-end libraries) (Fig. S4B.1). The TPM represent a measure of RNA abundance and hence, it provides an idea of gene expression levels in a particular sample. Interestingly, TPM distribution plots (percentiles) showed a distinct spike in expression levels towards percentile 98% for all our seven RNA-libraries (Fig. 4-4a). After the conservative filtration of

the 2% top expressed unigenes, I selected a subset of 1,437 unigenes with particular higher levels of expression (higher TPM values and PCA outliers) in *Oophaga* skin tissue (Fig. 4-4b). Homology *blast* analysis of these over expressed unigenes revealed that 76% (n=1,092) show significant similarities with annotated gene products and/or known protein domains (Fig. 4-5a) distributed mainly across amphibian (frogs) species (*X. tropicalis*=30%, *X. laevis*=14% and *Rana catesbiana*=4%) (Fig. 4-5b). Based on GO mapping level II, these unigenes were assigned to 22 BP, 19 CC and 11 MF (Fig. S4B.2). Multilevel GO term classification assigned the highly expressed unigenes to 16 BP, 5 CC and 9 MF (Fig. 4-5c). Within BP, 10% of unigenes were associated with oxidation-reduction processes, followed by response to stimulus (7.2%) and regulation of macromolecule metabolic processes (6.8%). Within CC, the integral components of the cell membrane were highly represented (35%), followed by protein complex (23%) and nuclear components (18%). Within MF, most of the highly expressed unigenes were assigned to metal ion binding activity (19%), followed by protein binding activity and transmembrane transporters (15%), oxidoreductase activity (14%) and structural molecule activity (10%) (Fig. 4-5c).

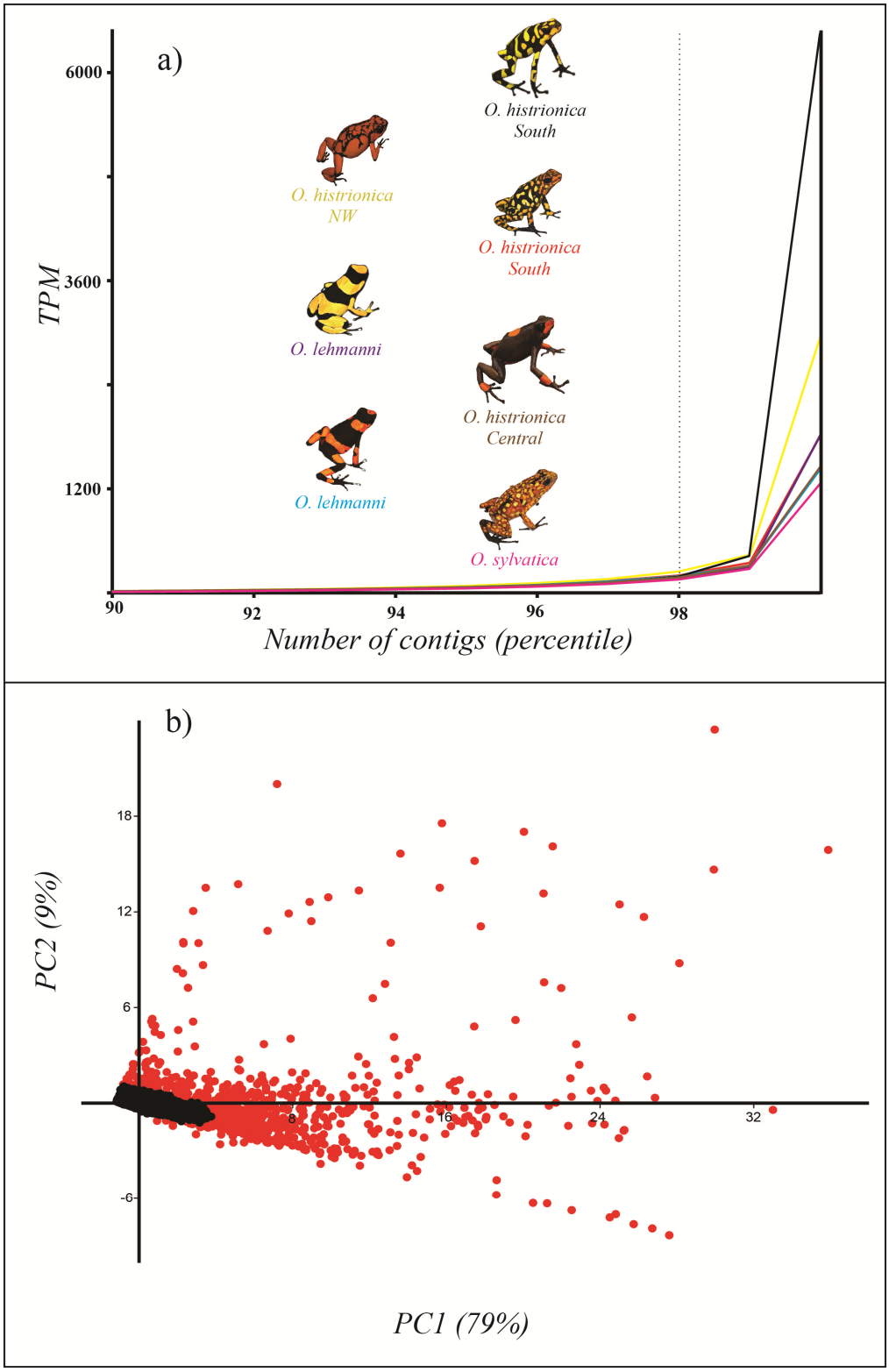


Figure 4-4. a) Percentile plot of the estimated transcripts per million (TPM) in the composite reference transcriptome as calculated based on raw RNA reads from individual libraries. Color names are equivalent to those in tendency lines. b) Principal component analyses (PCA) plot of the reference contigs dataset using TPMs values as independent variables. Red dots represent the selected over expressed unigenes (2% of the total contigs; n= 1,437) while dark dots represent the non-over-expressed contigs (n=30,061)

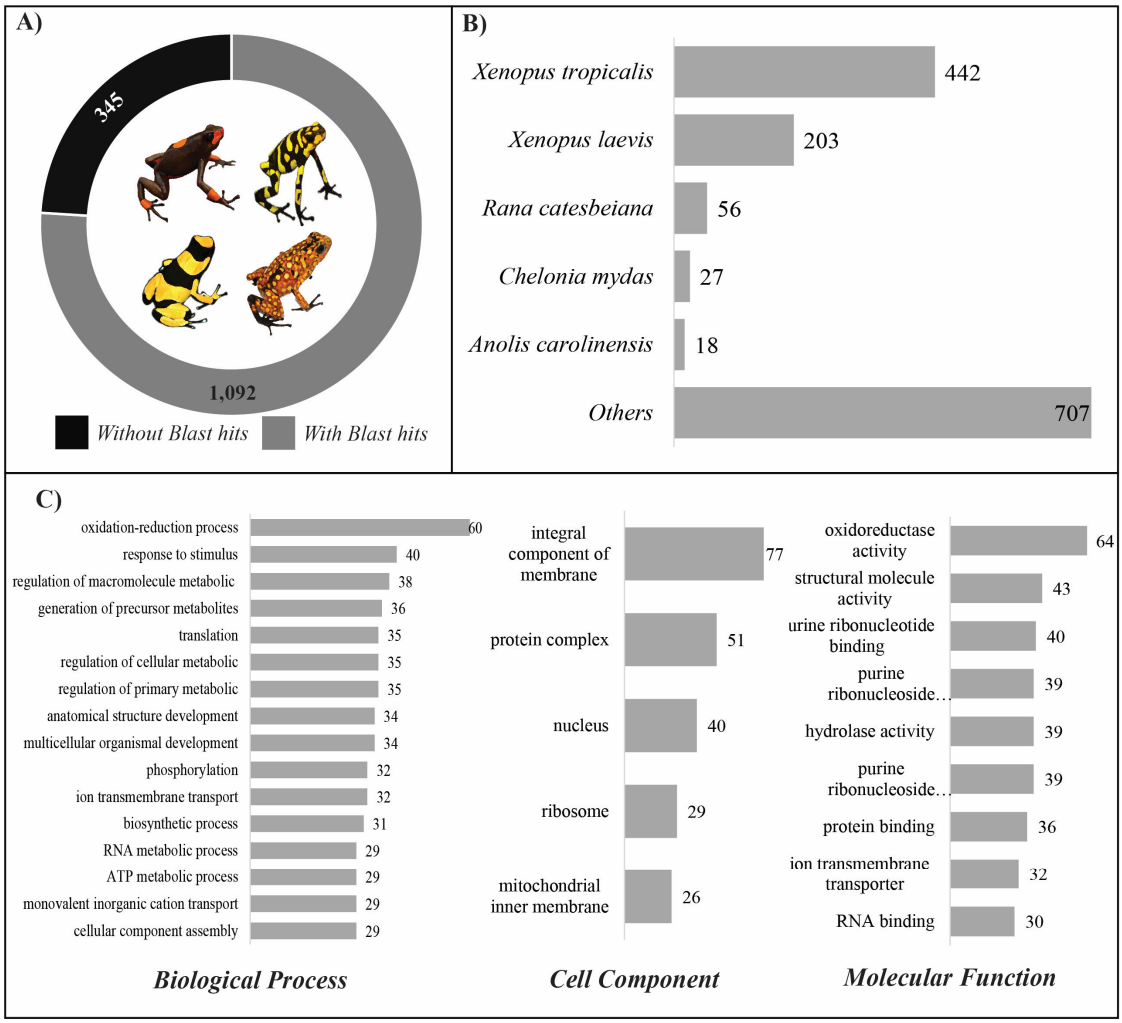


Figure 4-5. a) Pie chart representing the total number of over expressed unigenes and the proportion of contigs with significant *Blast* hits ($E<1.0E^{-5}$). b) Blast hit species distribution of over expressed unigenes. c) Gene ontology (GO) categories distribution (multilevel) for the annotated over expressed unigenes.

4.3.3 Phylogeny and genetic distances

The implemented workflow for SNC discovery allowed the identification of 330 orthologous markers all present in each individual transcriptome. The final dataset after manual editing yielded an alignment of 283,881 bp with 672 segregating sites (S), 413 parsimony informative sites and a nucleotide diversity (Pi) of 0.0017. Pair-wise genetic distances between individuals ranged from 0.07-0.26%. The greatest genetic distances are those observed between *O. sylvatica* and all other species (0.19-0.26%) while the lowest ones were observed between individuals from the same species (paired vs single-end transcriptomes) (Table 4-2). ML and BI phylogenetic reconstructions using the best model of nucleotide substitution (TN93+G) yielded trees with identical topologies and both resolved the phylogenetic relationship among these *Oophaga* lineages with a strong node support (bootstrap BS=100%; posterior probability PP>0.92). As shown in Figure 4-6, *O. sylvatica* represents an independent lineage with the longest branch length (i.e. larger amount of change) and closely related to *O. lehmanni* individuals. A second cluster (upper cluster in the tree, Fig. 4-6) included all *O. histrionica* individuals only. Within this nominal species, individuals were grouped in two distinct clusters that correspond to the lineages I-III and IV described in Chapter 3 (Posso-Terranova & Andrés, 2016b). ML phylogenetic reconstruction based on the 4,237 of putative orthologs (dataset #2) identified by reciprocal BLAST gave the same topology.

Table 4-2. Genetic distance estimate between individuals (TN93+G) based on 330 SCN markers. Standard errors were calculated by using the bootstrap method (n=999) and are shown above the diagonal. Minimum and maximum values are shown in bold letters.

	***O. histrionica* North**	***O. histrionica* Central***	***O. histrionica* South**	***O. histrionica* South***	***O. lehmanni***	***O. lehmanni****	***O. sylvatica****
***O. histrionica* North**		0.0000	0.0001	0.0001	0.0001	0.0001	0.0001
O. histrionica* Central	0.0010		0.0001	0.0001	0.0001	0.0001	0.0001
***O. histrionica* South**	0.0012	0.0012		0.0001	0.0001	0.0001	0.0001
O. histrionica* South	0.0015	0.0015	0.0009		0.0001	0.0001	0.0001
O. lehmanni	0.0019	0.0019	0.0015	0.0018		0.0000	0.0001
O. lehmanni*	0.0019	0.0019	0.0016	0.0018	**0.0007**		0.0001
O. sylvatica*	0.0024	0.0024	0.0023	**0.0026**	0.0019	0.0019	

*. Single-end libraries

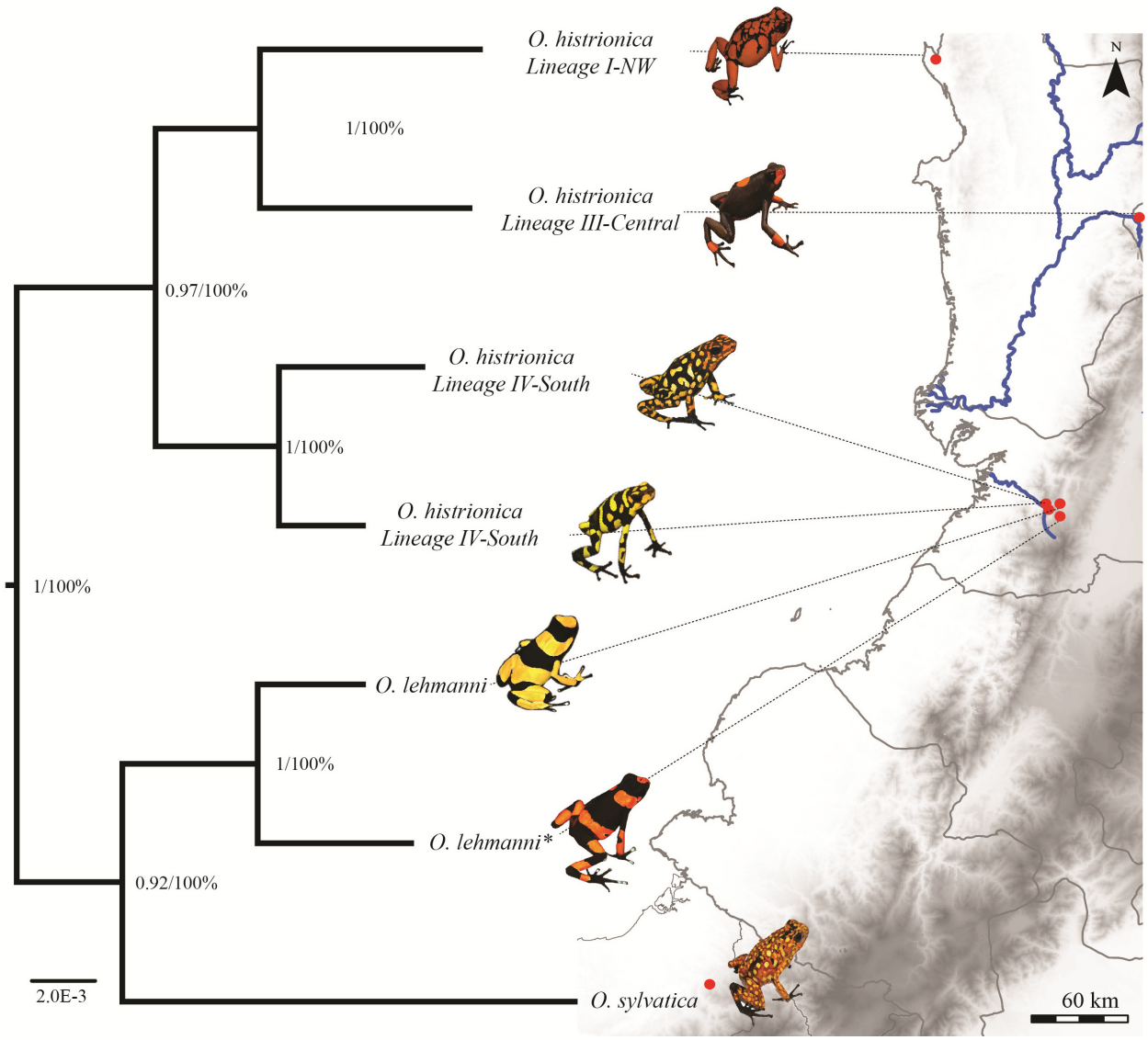


Figure 4-6. Maximum clade credibility (MCC) based on the analysis of 330 orthologous markers (dataset #1) including the seven individuals of the *O. histrionica* complex evaluated in this study. Numbers in nodes represent Bayesian posterior (BP) and maximum likelihood (ML) probabilities. A ML phylogenetic reconstruction based on 4,237 putative orthologs (dataset #2, see methods) produced a tree with the same topology. Species names in the figure correspond to the current taxonomic classification (*sensu lato* Myers & Daly, 1976; Funkhouser, 1956). *O. histrionica* lineages are those described in Chapter 3.

4.4 Discussion

4.4.1 Expression profiles reveal genes and mechanisms potentially related to alkaloid transportation and resistance to auto-toxicity

Because the alkaloids present in the skin of *Oophaga* frogs are very similar to those present in plants (Daly *et al.*, 1994; Daly *et al.*, 2002; Daly *et al.*, 2005; Daly *et al.*, 2009), it is reasonable to assume that the transportation and accumulation of alkaloids in these frogs may be carried out by homolog or similar systems to those described in plants were alkaloids are transferred from source to sink organs (Hashimoto & Yamada, 1994; Hartmann, 1999). If so, at least, two alternative membrane mechanisms may help to explain the transportation and storage of unmodified diet-derived alkaloid by the specialized cells of skin secretory glands. First, alkaloids in a lipophilic state may freely pass through the cell membranes by simple diffusion and accumulate in acidic secretory lysosomes if they become protonated to form hydrophilic cations. This ion-trap mechanism is not energy-dependent, and does not necessarily require the expression of any transporters. On the contrary, the transportation of alkaloids may be managed by proton-antiport carrier systems in an energy-requiring manner (Otani *et al.*, 2005; Shitan & Yazaki, 2007; Carqueijeiro *et al.*, 2013). Under this alternative model, diet derived alkaloids may be taken up by ABC transporters (Sakai *et al.*, 2002; Shitan *et al.*, 2003) and may accumulate in the secretory lysosomes of the skin gland cells by a cation exchanger antiporter (CAX) system, dependent on the pH gradient generated by "vacuolar" type ATPases ($V\text{-}H^+$-ATPases) and/or pyrophosphatases ($V\text{-}H^+$-PPases). To find some support for these two potential models that might explain the accumulation and posterior secretion of dietary alkaloids, I investigated the common transcriptional profile of all skin samples. In each lineage, the transcript abundance (TPM) analysis revealed that at least 12 homologs of different type II cation

exchanger proteins (CAX) are highly expressed in the skin of all the *Oophaga* lineages studied here (Table 4-3). A result suggesting the existence of an active membrane transportation-accumulation mechanism of alkaloids in Harlequin poison frogs. Consistent with this hypothesis, I also found putative ABC transporters and at least one V-H^+-ATPase on the skin transcriptome. All together, these results suggest that active lysosomal exocytosis may play a key role in the secretion of alkaloids in these frogs.

The accumulation of toxic compounds implies that organisms must avoid self-intoxication (auto-resistance). While the membrane transportation mechanisms described above reduce auto-toxicity by compartmentalizing the sequestered alkaloids, other non-alternative excluding mechanisms are likely to contribute to auto-toxicity resistance. In Harlequin poison frogs, the major toxic components present in skin tissue are histrionicoxins (HTX) (Daly *et al.*, 2005). These alkaloids are known to cause temporary paralysis or even death by inactivating or blocking voltage-gate ion channels (Daly *et al.*, 1985). Resistance to this type of cytotoxic compounds usually arises through either an increased expression level of P450 enzymes (CYPs) that metabolize the toxin, or through target insensitivity via mutations that reduce the toxin's ability to bind the ion channel itself (Chaudhary & Willett, 2006). My analyses revealed that five CYP homologs are amongst the most highly expressed genes in the skin of *Oophaga* lineages (Table 4-3). A result that suggests that oxidative biotransformation of lipophilic alkaloids to hydrophilic compounds (Sigel *et al.*, 2007) is an important auto resistance mechanism in these frogs. Additionally, I also found seven genes encoding voltage-gated ion channel proteins (VGIC). One of which, encodes the beta-3 subunit of the sodium channel in which six types of amino acid replacements in the inner pore (Nav1.4) are known to increase resistance to toxic alkaloids in other Dendrobatid frogs (Tarvin *et al.*, 2016).

4.4.2 Molecular bases of warning signals: candidate genes related to melanin-based colouration

The different lineages of the *O. histrionica* complex species studied here display a wide variety of warning colours. However, they also share a black background colouration (Fig. 4-1) which is possibly related to an increased aversion learning of the aposematic signal relative to one of solid bright colours alone (Bowdish & Bultman, 1993). In anurans, colouration relates to both the structure and pigment composition of the dermal chromatophore units in which three types of pigment cells (xantophores, iridiophores, and melanophores) are laid one on another (Bagnara *et al.*, 1968). In dark skin areas, the distal fingers of the melanophores are filled with melanin granules (melanosomes) obscuring the pigment of the other cells of the unit. Thus, genes involve in the amount, size and distribution of the melanosomes should play a significant role in the colouration pattern of these frogs. Accordingly, the comparative transcriptome analysis revealed the presence of five highly expressed genes potentially related with melanin-based colouration (Table 4-3).

In one of the few detailed studies of colouration in Anurans, Bagnara *et al.* (1979) suggest a common origin for of all the pigment granules found in the cells of the chromatophore: a primordial organelle (vesicle) derived from the rough endoplasmic reticulum (RER). According to this model, in the formation of melanosomes, the premelanosomes are derived from cisternae of the RER, which then fuse with vesicles derived from the Golgi complex containing tyrosinase enzymes (Bagnara & Taylor, 1970; Bagnara *et al.*, 1973; Bagnara *et al.*, 1978; Bagnara, 1982; Palumbo *et al.*, 1997). Two highly expressed genes identified here (tyrosinase precursors) (12213 and 15238, Table 4-3) may contribute to regulate this mechanism, which in turn might translate to differences in background colouration.

Interestingly, within the *histrionica* complex, I have found populations that are characterized by light-brown background colouration as oppose to black (Fig. 3-2 in chapter 3 and Fig. 5-1 in chapter 5). One tantalizing possibility is that this phenotypic difference is indeed associated with differences in the production of cell pigments (melanin) and the formation of the melanosomes (the melanin-containing organelles) –see Chapter 5–.

Table 4-3. The table shows the identified over-expressed transcripts with potential important functions in the alkaloid sequestration system, auto-resistance to toxic compounds and variation in colouration in Harlequin poison frogs.

Transcript ID	Blast description	Length (bp)	Mean TPM	e-Value	Sim. mean	#GO terms	GO Names list	Potential role
contig_20619	cation-transporting ATPase 13A4-like	417	124.175	2.00E-47	81%	-		Alkaloid transportation
contig_21813	cation-transporting ATPase 13A5-like	523	85.039	9.00E-67	85%	-		
contig_7_AC	ATPase, Ca++ transporting, cardiac muscle, fast twitch 1	1890	459.739	0.00E+00	98%	7	F:ATP binding; F:metal ion binding; P:metabolic process; C:integral component of membrane; P:calcium ion transmembrane transport; F:calcium-transporting ATPase activity; C:sarcoplasmic reticulum membrane	
contig_1361	transmembrane 176A	1097	291.686	3.10E-06	47%	-		
contig_4021	transmembrane 215-like	484	240.418	1.00E-30	70%	-		
contig_2014	transmembrane 38A	1421	64.724	1.50E-126	94%	5	C:nuclear membrane; P:potassium ion transmembrane transport; C:integral component of membrane; F:potassium channel activity; C:sarcoplasmic reticulum membrane	
contig_9808	transmembrane 50A-like	723	182.707	1.20E-77	91%	-		
contig_3777	transmembrane BAX inhibitor motif containing 6	782	219.686	1.80E-83	92%	2	P:negative regulation of apoptotic process; C:integral component of membrane	
contig_10175	transmembrane emp24-like trafficking 10 precursor	783	224.969	3.90E-83	90%	2	C:integral component of membrane; P:transport	
contig_11539	transmembrane LOC401397	359	243.536	3.00E-06	83%	-		
contig_494	transmembrane trafficking (TMP21)	809	224.745	1.20E-95	93%	2	C:integral component of membrane; P:transport	
contig_5181	transport Sec61 subunit gamma-like isoform X1	1054	199.671	1.20E-20	100%	-		
contig_1632	FXYD domain containing ion transport regulator 1 precursor	1009	179.973	1.20E-22	91%	3	F:ion channel activity; P:ion transmembrane transport; C:integral component of membrane	Alkaloid auto-resistance: target-site insensitivity
contig_15371	FXYD domain-containing ion transport regulator 3 precursor	743	448.948	1.00E-18	89%	-		

contig_3820	calcium channel, voltage-dependent, gamma subunit 1	524	36.551	8.00E-39	98%	4	C:voltage-gated calcium channel complex; P:regulation of ion transmembrane transport; P:calcium ion transmembrane transport; F:voltage-gated calcium channel activity	
contig_1037	sodium channel subunit beta-1	1316	40.39	5.50E-67	78%	-		
contig_376	voltage-dependent anion channel 2	388	414.009	5.20E-43	98%	6	F:voltage-gated anion channel activity; P:regulation of anion transmembrane transport; P:transmembrane transport; C:integral component of membrane; C:mitochondrial outer membrane; P:anion transport	
contig_907	voltage-dependent anion channel 2	990	271.662	9.90E-86	100%	6	F:voltage-gated anion channel activity; P:regulation of anion transmembrane transport; P:transmembrane transport; C:integral component of membrane; C:mitochondrial outer membrane; P:anion transport	
contig_12004	voltage-gated hydrogen channel 1	1252	161.114	7.10E-53	88%	6	C:integral component of plasma membrane; P:proton transport; P:cellular response to zinc ion; F:voltage-gated proton channel activity; P:regulation of ion transmembrane transport; P:cellular response to pH	
contig_14682	cytochrome P450	591	145.905	3.70E-37	70%	5	F:arachidonic acid epoxygenase activity; P:epoxygenase P450 pathway; F:metal ion binding; C:intracellular membrane-bounded organelle; F:steroid hydroxylase activity	Alkaloid auto-resistance: metabolic detoxification
contig_13814	cytochrome P450 2F2-like	537	283	6.20E-53	72%	-		
contig_7532	cytochrome P450 2F2-like	1201	93.48	3.30E-76	70%	-		
contig_1039	cytochrome P450 2K1-like	429	301.328	1.70E-57	87%	-		
contig_12974	Cytochrome P450, family 2, subfamily f, polypeptide 2	419	271.474	4.20E-37	70%	9	F:arachidonic acid epoxygenase activity; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen; P:trichloroethylene metabolic process; P:epoxygenase P450 pathway; F:metal ion binding; C:intracellular membrane-bounded organelle; F:steroid hydroxylase activity; P:response to toxic substance; P:naphthalene metabolic process	
contig_12213	premelanosome precursor	1454	189.702	4.50E-158	78%	1	C:membrane	Color ation

contig_15238	tyrosinase-related 1 precursor	2041	77.418	0.00E+00	90%	6	C:melanosome membrane; P:melanin biosynthetic process; F:copper ion binding; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, another compound as one donor, and incorporation of one atom of oxygen; C:integral component of membrane; P:oxidation-reduction process
contig_26937	melanocortin 1 receptor	2181	65.634	3.50E-139	90%	-	
contig_14447	melanoregulin-like [Xenopus (Silurana) tropicalis]	1131	87.066	6.20E-66	68%	-	
contig_12696	melanophilin [Xenopus (Silurana) tropicalis]	998	196.305	7.40E-161	75.50%	-	

There are other molecular and cellular mechanisms that might be associated to the difference between light and dark background colouration. In lower vertebrates, dark hues are known to be produced by the interaction between high levels of melanocyte-stimulating hormone (α-MSH) and several variants of its transmembrane receptor (*MC1R*) through the dispersion of melanosomes within the melanophore (by increasing cAMP intracellular levels) (Sugimoto, 2002; Logan *et al.*, 2006). Thus, it is possible that structural or expression differences in the MC1R might contribute to dark phenotypes in *Oophaga* frogs. While there is a strong evidence that different mutations at *MC1R* cause either light or dark phenotypes in many mammals, birds and reptiles (Everts *et al.*, 2000; Gross *et al.*, 2009; Gangoso *et al.*, 2011; Baião & Parker, 2012; Corso *et al.*, 2012), the only two studies conducted in frogs are inconclusive (Herczeg *et al.*, 2010; Chikako, 2012). My results indicated high transcription levels of this G-receptor in the skin of all the studied lineages. A detailed inspection of the coding sequences recovered for this gene revealed the presence of different length isoforms, making *MC1R* a promising candidate gene candidate to explain the differences in background colouration in poison frogs (see Chapter 5). Finally, the comparative study unraveled another two highly expressed genes (14447-melanoregulin- and 12696-melanophilin) that may also contribute to dark hues. In this case, the predicted products of these genes are key proteins that mediate the melanosome transport and distribution in epidermal cells through the formation of a tripartite protein complex (Ohbayashi *et al.*, 2012). The disruption of the transport protein complex results in pigmentary dilution and lighter phenotypes by the clustering of melanosomes around the nucleus (Van Gele *et al.*, 2009)

4.4.3 Phylogenomic insights from SCN markers

Despite the increased amount of genomic data provided by NGS technologies and

improved computational tools, the development of informative markers from transcriptomic data with the potential to resolve phylogenetic relationships remains still underexplored. Only a few studies have reported success in developing large sets of SCN markers by using transcriptomics data, however, these investigations required the use of sophisticated bioinformatics pipelines and the access to relatively advanced computing facilities (reviewed in Chamala *et al*.,2015). In this study, I report a workflow that allowed the identification of hundreds of SNC orthologous markers (dataset #1, n=330) without the need of advanced bioinformatics training and using limited computer resources (Fig. 4-2). This approach takes advantage of independently combining several bioinformatics tools already available for the detection, comparison, filtering and alignment of highly conserved eukaryotes orthologous. It also uses an expanded database of eukaryotic orthologous (Parra *et al.*, 2007) without focusing only in those available for angiosperm plants (i.e. MarkerMiner)(Van Bel *et al.*, 2012).

I am well aware that the SNC loci detected here may correspond to highly conserved constitutive genes (housekeeping) that are required for the maintenance of basic cellular function. Then, these markers may be of limited utility to resolve the phylogenetic relationship of recently and rapidly diverging evolutionary lineages. Despite this, the identified SCN markers allowed me to generate a robust dataset of high-quality single nucleotide (SNP) and indel polymorphisms to compare the seven individual poison frogs. Because of the low number of individuals in this study, I cannot rule out that a portion of the SNPs is represented by individual rather than population- specific polymorphisms. Nevertheless, considering the high number of SNPs, this analysis resolved the relationship between these *Oophaga* frogs with a strong statistical support (Fig. 4-6) and provide a phylogeny that is highly concordant with previous studies (Medina *et al.*, 2013; Posso-Terranova & Andrés, 2016a; Posso-Terranova &

Andrés, 2016b). In order to confirm the robustness of this approach, I also performed a phylogenetic reconstruction based on a set of 4,237 putative orthologs (dataset #2), which produced a tree with exactly the same topology. Here, I provide evidence from genome-wide transcriptomic data indicating that within the currently known *O. histrionica* species (Myers & Daly, 1976), there are at least three different independent evolving lineages comparable in divergence with the other two recognized species in the phylogeny (*O. sylvatica* and *O. lehmanni*) (Posso-Terranova & Andrés, 2016b). Thus, the SNP analysis of SCN orthologous markers proved to be powerful for studying evolutionary processes in groups with young evolutionary histories and shallow phylogenies, even on a minimal number of samples.

Overall, this study demonstrates the utility of using RNA-sequencing with non-model organisms to identify loci that may be of adaptive importance. Altogether, these data enabled me to provide the first global study of *Oophaga* poison frogs transcriptomes and allowed me to proposed a set of candidate genes underlying important functions in the alkaloid sequestration system, auto-resistance to toxic compounds and variation in colouration and patterns. The skin transcriptomes reported here enhance our current knowledge of tissue-specific gene expression in Harlequin poison frogs and constitute a valuable genomic tool for future biochemical, molecular and functional studies.

CHAPTER 5

STRUCTURAL BASIS OF BACKGROUND COLOURATION AND ITS RELATION WITH THE MELANOCORTIN 1 RECEPTOR (MC1R) IN HARLEQUIN POISON FROGS

5.1 Introduction

One of the main goals of evolutionary genetics is to reveal the links between genotype, phenotype and fitness. colouration variation in vertebrates are promising systems to explore such connections. Within species, some vertebrates show a striking range of colour variation including polymorphism (Crothers *et al.*, 2011), sexual dimorphism (Maan & Cummings, 2009), behaviourally induced colour displays (Bagnara & Hadley, 1969) and/or rapid plastic colour changes (Selz *et al.*, 2007). Many reasons explain why differential colouration patterns within species have evolved: in some species, colour signals enable individuals to communicate information about their sexual status or their ability to defend itself while in others, colouration help individuals to avoid predation or regulate their temperature (Poulton, 1890; Auber, 1957; Searle, 1968; Endler, 1977; Endler, 1980; Hoffman & Blouin, 2000; Griffith *et al.*, 2006; Logan *et al.*, 2006; Gratten *et al.*, 2007; Endler, 2012). Not surprisingly, numerous studies have focused on identifying the selective mechanisms and fitness correlates of colouration (Harvey *et al.*, 1982; Forsman & Shine, 1995; Hagman & Forsman, 2003; Allison & Funk, 2009). Similarly, because vertebrate pigmentation has served as a model system to learn about gene interactions (Silvers, 1979), more than a hundred genes that affect animal colour and patterning have been identified (Hoekstra, 2006; Protas & Patel, 2008; Mundy, 2009). Despite this relative abundance of either genetic or ecological/behavioural studies on colouration, even today, only few studies have been able to clearly demonstrate the links between genotype and colouration traits under selection in natural populations. For

example, in poikilotherms (particularly amphibians), most of the research efforts far have focused on the cellular structural and histological components of colour variation (Bagnara *et al.*, 1968; Bagnara & Hadley, 1969; Bagnara *et al.*, 1973) and only recently, progress has been made towards understanding how the astonishing anuran's variation in colouration is linked to their ecological, behavioural and evolutionary aspects (Hoffman & Blouin, 2000; Herczeg *et al.*, 2010; Chikako, 2012; Rudh & Qvarnström, 2013; Rojas, 2016).

The melanocortin-1 receptor (MC1R) is known to play a key role in the colour of a broad range of species. In homeotherms, this G-protein coupled receptor is a key switch in the biochemical pathway that leads to the production of different melanin pigments (dark eumelanin; light/yellow/red phaeomelanin) and hence, it has been associated with intra and interspecific colour variation in wild (Rosenblum *et al.*, 2004; Dun *et al.*, 2007; Gangoso *et al.*, 2011; Baião & Parker, 2012; Domingues *et al.*, 2012). Although anuran's pigment cell morphology differs from that of mammals and birds, the function of MC1R is remarkably conserved (Rees, 2000; Rosenblum *et al.*, 2004). As such, in both fishes and reptiles, MC1R activity is related to the amount of melanin produced and regulates melanosome size and dispersion within melanophores (Sugimoto, 2002; Rosenblum *et al.*, 2010). The poison frogs of the *Oophaga histrionica* complex (Myers & Daly, 1976; Lötters, 1992; Lötters *et al.*, 1999; Medina *et al.*, 2013) are an excellent model system to examine the potential role of MC1R in colour variation in anurans. While all the members of this complex advertise their toxicity to predators by using warning bright colours, there is striking variation in the colouration patterns among lineages including contrasting differences between dark-black and light-brown background colouration (Fig. 5-1). Here, I evaluated the variation in background colouration at three hierarchical levels. First, I document the geographic pattern variation among the closely

related lineages of the *O. histrionica* complex. Second, I examine if the difference between light-brown and dark-black correlate with histological and structural characteristics previously documented in other Anurans. Third, I compare intra- and interspecific variation at MC1R with colour variation as a first step toward identifying the genetic basis of aposematic colour variation in natural populations of *Oophaga* poison frogs. My results showed that dark-black background phenotypes in *Oophaga* poison frogs is explained by an increased melanosome size and the absence of a "typical" dermal chromatophore unit (Bagnara *et al.*, 1968). Furthermore, I showed that a dark-black phenotype has arisen independently in geographic isolated *Oophaga* lineages through different mutations at MC1R that, in this phenotype, is associated with highly truncated MC1R receptors. This study offers the first insights on the structural and molecular underpinnings of colour variation in Dendrobatid poison frogs and exemplify the potential effects of MC1R in this important adaptive trait.

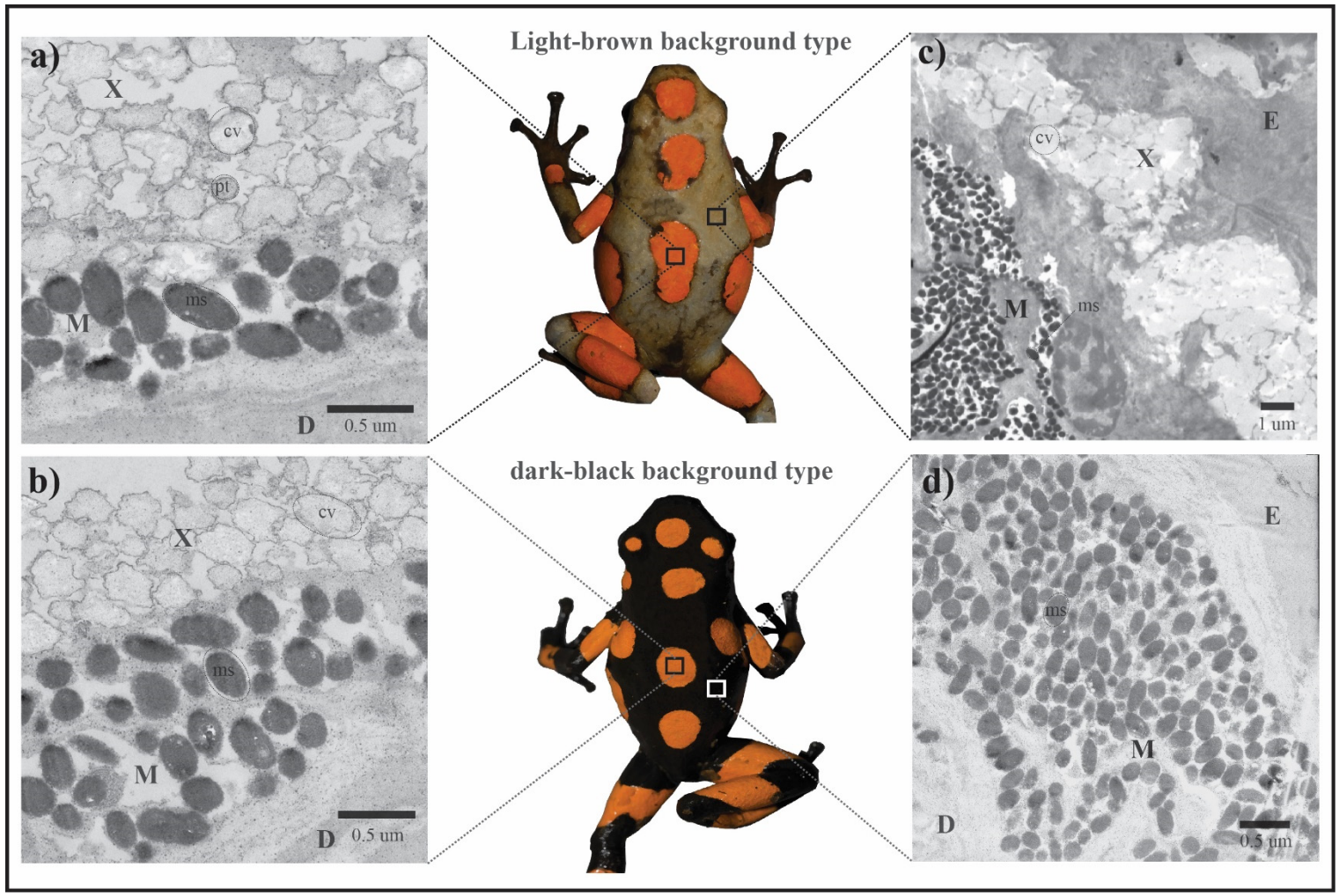


Figure 5-1. Individuals of *O. histrionica Lineage II-NE (*upper) and *O. histrionica Lineage I-NW* (lower; see Chapter 3) with contrasting background colouration. a-b) Electronic microphotographs of ultrathin skin cuts from dorsal coloured spots. c) Skin cuts from dorsal light-brown and (d) dark-black background colouration. X=xantophores, M=melanophores, D=dermis, E=epidermis, cv=carotenoid vesicle, ms=melanosome, pt=pterinosome.

5.2 Materials and Methods.

5.2.1 Variation in dorsal background colouration

To study the variation on background colouration within the *O. histrionica* complex, I obtained colouration-intensity data from 215 individuals across 15 geographic localities (Fig. 5-2), as follows: for each individual, ventral and dorsal raw digital images (Nikkon Electronic Format: NEF, Nikkon D610) were taken using an Spectralon grey reflectance standard (Labsphere, Congleton, UK) as a background. Raw images were then linearized with respect to light intensity (Stevens *et al.*, 2007) and stored in TIFF format (lossless compression). Images were used to collect data on 12 variables related with dorsal and ventral colouration intensity (Supplementary Table S5B.1) using ImageJ v. 1.48 (Sheffield, 2007). The resulting data matrix was then summarized by implementing a principal component analysis (PCA). To assess colouration similarity in a multivariate framework, I performed a hierarchical clustering routine (method= average; distance=Gower; cluster support= 1000 bootstrap) as implemented in PAST v. 3.08 (Hammer *et al.*, 2001) followed by a permutational analysis of variance (PERMANOVA, permutations=9999) to test the null hypothesis that the centroids of the resulting clusters are equivalent for all groups. As an indicator of the level of individual chromatic contrast, I measured the brightness difference between the dorsal coloured and background brightness indexes (estimated as described in Table S5B.1). Greater values of brightness difference indicate a higher light/dark contrast between a colour and the background where it occurs. Pairwise comparisons between group levels were performed using the '*pairwise.t.test*' function of the R-package (Ihaka & Gentleman, 1996).

5.2.2 Structural bases of colouration

Dorsal skin samples were obtained from two adult individuals from two different *Oophaga* lineages with contrasting background colouration: Linage I, dark-black, and Linage II, light brown. (Fig. 5-1). Specimens were euthanized in the field with benzocaine gel at 5% and tissue samples from both, reddish-orange coloured spots (Fig. 5-1a, b), and light or dark black background (Fig. 5-1c, d) were taken and stored in 10% neutral buffered formalin (NBF) until further analysis (Puchtler & Meloan, 1985). Histological analysis were carried out by transmission electron (TEM) and bright-field (BFM) microscopy (Bagnara *et al.*, 1968; Bagnara & Taylor, 1970; Frost & Robinson, 1984). Skin samples were fixed in 3% glutaraldehyde buffer with 0.1 M cacodylate at pH 7.4 for 4 hr at 4ºC. Fixed samples were then soaked for 4h in 2% osmium tetroxide, rinsed and dehydrated in a graded ethanol series. Dehydrated samples were embedded in Epon 812 and cut on a Reichert-Jung Ultracut-E ultramicrotome in either semi-thin (1-10µm) or ultrathin (60-100nm) sections using a diamond blade. Semi-thin sections suitable for BFM were stained with toluidine blue (1% in 70% ethanol) while ultrathin sections suitable for TEM (Philips CM 10) were first collected on Formvar-coated carbon-stabilized grids, and then stained for 30 minutes with 2% uranyl acetate. Analyses of melanosome morphology were performed as described by (Alaluf *et al.*, 2001). Briefly, monochrome TEM images with melanosomes sparsely distributed were copied from the original microphotograph to a second image with a plain white background. The resulting images were then analysed with ImageJ v. 1.48 (Sheffield, 2007). After scaling, the area and the circularity (short:long axis ratio) of individual melanosomes were measured (n ≥25 melanosomes/section). The density of melanosomes was calculated as the proportion of a total area occupied by dermal melanosomes on each individual section. Variation in melanosome size, circularity and density of

melanosomes from different skin portion was modelled using a one-way ANOVA (Alaluf *et al.*, 2002). Post-hoc pairwise comparisons with corrections for multiple testing were performed using the '*pairwise.t.test*' function included in the R-package (Ihaka & Gentleman, 1996).

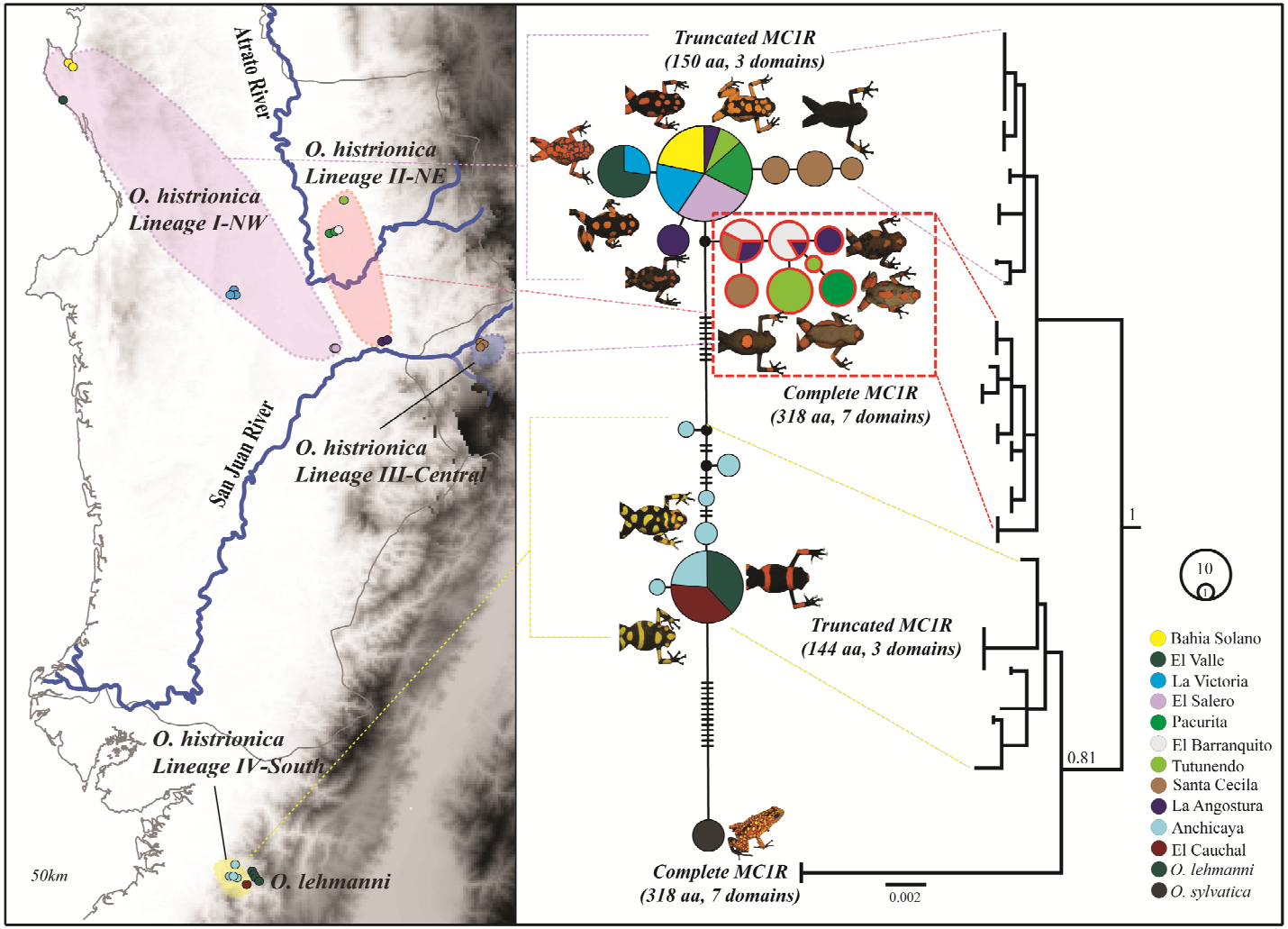


Figure 5-2. Map showing the potential distribution of five *Oophaga* lineages evaluated in this study. Coloured dots within the map represent sampling localities. A network representing the relationship between different MC1R haplotypes and a corresponding phylogenetic MC1R tree are shown in the middle and right side of the figure. The tree was condensed at 50% bootstrap support (or 0.5 posterior probability; PP). Number in nodes correspond to PP and only values >0.8 are shown in the tree.

5.2.3 Isolation and PCR amplification of MC1R

Tissue samples from *O. histrionica* lineages I (n=4) and II (n=4) were obtained by toe-clipping. Total genomic DNA was isolated using MasterPure^{TM} DNA purification kit (Epicentre, Madison, WI, USA) according to the manufacture's protocol. Using PrimerSelect (Lasergene, DNASTAR), I designed primers against conserved amino acid residues across a wide range of vertebrates. Initially the PCR primers (MC1F/R) amplified a 570-bp fragment of the MC1R (Table 5-1). Then, using the genome walk strategy described previously (Deng *et al.*, 2010), I extended the gene sequence in both the 3' and 5'gene directions using degenerate and gene-specific primers . The open reading frame of the resulting fragments were determined using EditSeq (Lasergene, DNASTAR). The predicted amino acid structure were then compared with that of well characterized MC1R vertebrate proteins, including two frog species (*Rana temporaria* and *Xenopus laevis*). Using new sets of primers (MC1R-13-1F/R and MC1R-13-2-F/R; Table 5-1) designed to amplify the full-length MC1R, I surveyed the nucleotide variation in a sample of 62 individuals from the same populations from which I collected phenotypic data (*O. histrionica* Lineage I*: :* n_{ind}= 18; Lineage II*: :* n_{ind}= 22; Lineage III*:* n_{ind}= 8; Lineage IV=*:* n_{ind}= 10; *O. lehmanni:* n_{ind}= 4). Additionally, I included two *O. sylvatica* individuals in the analyses as an outgroup species. PCR amplification, Sanger sequencing and sequence editing protocols were performed as previously described (Hauswaldt *et al.*, 2011; Brusa *et al.*, 2013; Medina *et al.*, 2013). As background colouration differences were not highly contrasting (as evaluated by human eye) among *O. histrionica* lineage III individuals (Santa Cecilia populations, Fig. 5-2), to assess colouration similarity among the genotyped individuals, I performed a permutational analysis of variance (PERMANOVA, permutations=9999) followed by a PCA of the previously described colour variables (Table S5B.2).

Table 5-1. Description of PCR and sequencing primers used in this study.

Name	Sequence (5' - 3')	Amplicon size	Annealing temperature (°C)
MC1F	GAACCCATGGTCAAAATGATG	570	55
MC1R	AAGGCGTAGATAATAGGGTCAA	570	55
MC1R_13_1F	GAATCTTTCGGGGAGACGTT	1127	65
MC1R_13_1R	TGGGAGATGGAGAGGACTTG	1127	65
MC1R_13_2F	GCAGGGTGAAAGGGAATCTT	1144	60
MC1R_13_2R	TTCTTGGGAGATGGAGAGGA	1144	60

5.2.4 Evolutionary analyses

Haplotype sequences from heterozygous individuals were estimated from the population data using PHASE v. 2.1.1 (Stephens *et al.*, 2001). Pairs of SNPs whose phase could not be statistically inferred with a posterior probability value ≥ 0.9 were scored as missing data. Phased MC1R haplotypes (*i.e.* alleles) were aligned using the *Clustal-W* algorithm (Thompson *et al.*, 1994). The resulting alignment was then used to estimate the best model of nucleotide substitution using the sample size-corrected Akaike Information Criterion (AICc) (Posada & Buckley, 2004) as implemented in jModelTest v2.1.4 (Posada, 2008).

MC1R genealogies were constructed using maximum likelihood (ML) and Bayesian inference (BI). Using MEGA6 and the results derived from jModelTest, I obtained an ML tree using extensive (level 5) Subtree Pruning and Regrafting (SPR) heuristic searches. The relative support for each node on this tree was estimated by generating 1,000 bootstrap replicates. The BI-based genealogy was generated using BEAST v. 2.1.2 (Drummond *et al.*, 2012; Bouckaert *et al.*, 2014) with a relaxed molecular on clock and an uncorrelated log-normal (UCLN) model of molecular rate heterogeneity. Three chains were run for 10 million generations sampled every 1000 steps. Trees and log files from the three independent runs were combined using

LogCombiner 2.1.2. and resampled at a lower frequency (i.e. 400 steps or ¼) to reduce the number of final trees (Bouckaert *et al.*, 2014). For each estimated parameter, convergence was assessed using Tracer v. 1.6 (http://tree.bio.ed.ac.uk/software/tracer/) and effective samples sizes (ESS) were calculated to ensure adequate mixing (ESS>350, after 30% burn in). The posterior probability density of the combined tree and log files was summarized as a maximum clade credibility (MCC) tree using TreeAnnotator v. 2.1.2 (Bouckaert *et al.*, 2014). Relationships among MC1R alleles were also plotted onto a TCS haplotype network with a connection limit of 95% (Clement *et al.*, 2000) implemented in the program *PopART* (Population Analysis with reticulate trees) available at http://popart.otago.ac.nz/.

DNA sequence variation estimates were calculated using DNAsp v.5.10.01 (Librado & Rozas, 2009) and genetic distances were estimated using the best model of nucleotide substitution in MEGA6 (Tamura *et al.*, 2013). I scanned for signals of selection by performing Tajima's D (Tajima, 1989) and Fu and Li's D tests (Fu & Li, 1993). Because these tests of selection based on the frequency spectrum are likely to be influenced by the demographic history and the structure of the populations , I also performed codon based Z-tests and McDonald–Kreitman style tests of neutrality (McDonald & Kreitman, 1991) where I calculated neutrality indices (Rand & Kann, 1996)for sequences between truncated alleles sharing a common stop codon as well as between truncated and full length alleles. Under the assumption that silent mutations are neutral, according to the usual interpretation of the MK test, genes with NI ≥ 1 and a significant *G*-test represent genes with a significant excess of non-synonymous polymorphisms, whereas genes with NI < 1 and a significant *G*-test represent cases of positive selection. All tests were performed for the entire coding portion of the gene using the methods implemented in DNAsp v.5.10.01 (Librado & Rozas, 2009).

5.3 Results

5.3.1 Variation in dorsal colouration

The multivariate analyses of 12 colour-intensity measurements showed that the first two PCA axes explained 96.2% of the total variation and clustered all sampled populations in two groups with contrasting background colouration: dark-black and light-brown individuals (Fig. 5-3). Accordingly, the hierarchical clustering analysis indicated the same aggrupation with high statistical support (BS=100%, Fig. 5-3). A few individuals with intermediate phenotypes showed an scatter distribution along the clustering tree, however, PERMANOVA analyses confirmed the differentiation of the two PCA groups (dark-black and light-brown background individuals) with high statistical significance ($F=274.5$, $p=0.0001$). Pairwise T-test of chromatic-contrast (brightness difference) showed that dark-black phenotypes either from the northern (Lineage I and III, $\bar{x}=89.57\pm 1.86$) and southern distributions (Lineage IV *and O. lehmanni;* $\bar{x}=108.90\pm 2.48$) were significantly higher (higher contrast between colours) and statistically different to the group with light-brown background colouration (Lineage II*;* $\bar{x}=8.14\pm 4.99$)($P<0.01$) (Fig. 5-2).

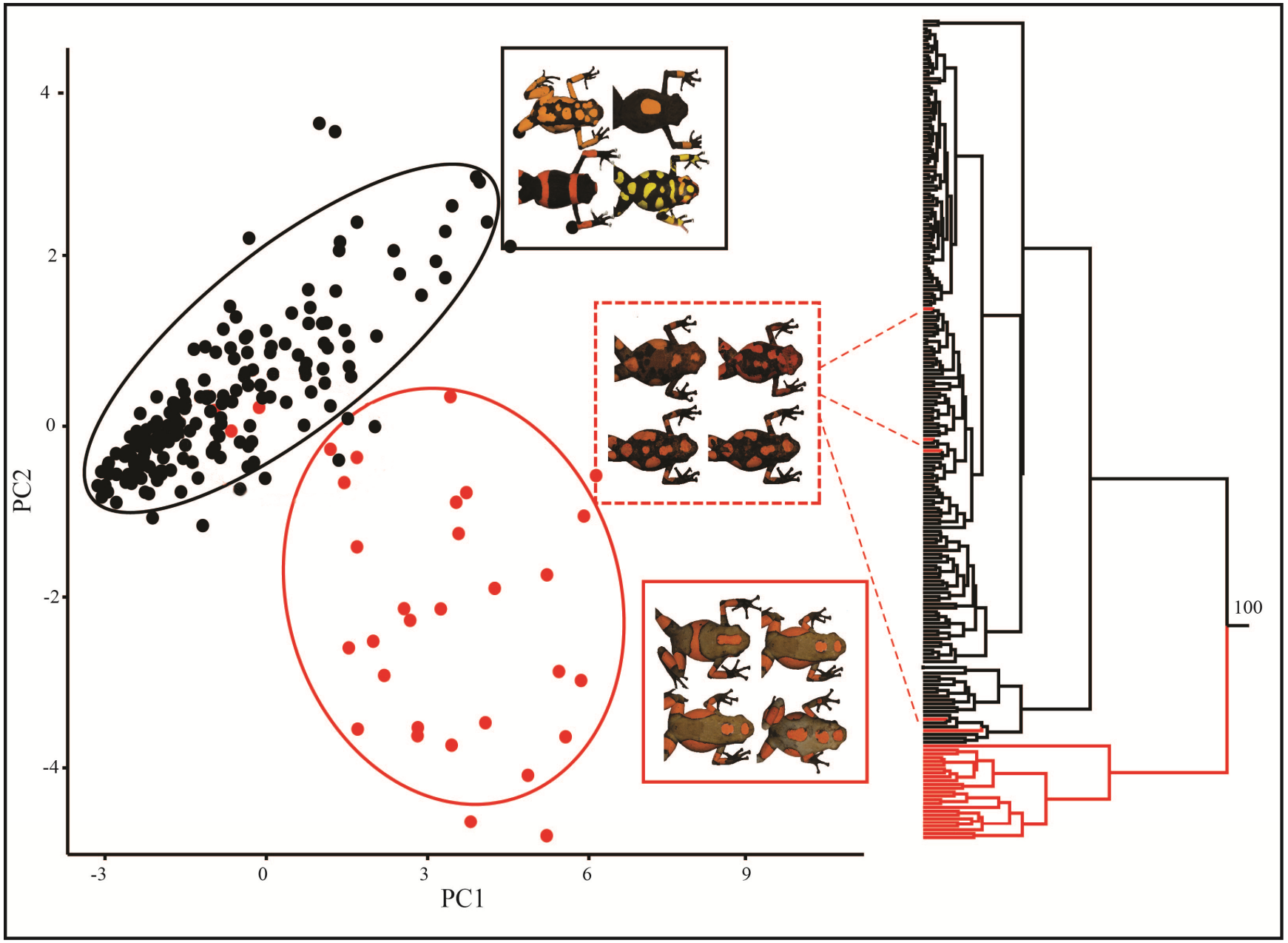

4
2
0
-2
-4
PC2
-3
0
3
6
9
PC1
100

Figure 5-3. Left: Two dimension clustering plot of individuals after a PCA analysis based on eight colour-related variables (PC1= 73%, PC2=23%). Solid ellipses include 95% of individuals. Red ellipse encloses individuals with light-brown background colouration while black encloses the dark-black ones. Dotted red lines indicates individuals with intermediate background colouration. Right: Hierarchical tree based on Euclidean distances of eight coloured-related variables indicating the clustering of light-brown and dark-black individuals. Node numbers represent bootstrap support (n=999) and only values >70% are shown.

5.3.2 Histology and colour variation

5.3.2.1 Coloured skin

Electron microphotographs of coloured skin portions from both types of individuals (light-brown and dark-black background) revealed the presence of two types of chromatophores (melanophores and xanthophores) (Fig. 5-1a and 5-1b). BFM also confirmed that xanthophores were closer to the epidermis while melanophores were located at a deepest layer towards the dermal region (collagen) (Fig. S5A.1). Within xanthophores, carotenoid-containing vesicles (cv) were of irregular shape and their size ranged from 0.1 to 0.3 um in both types of individuals. I found smaller (~0.1 um) and rounded organelles within the xanthophores that possibly corresponded to pteridine-containing pterinosomes (pt) (Fig. 5-1a), which has been reported in coloured skin portions of other amphibians (Frost & Robinson, 1984; Bagnara, 2003). Melanosomes (ms) within melanophores in coloured skin portions (Fig. 5-1a and 5-1b) were randomly arranged, typically rounded to ellipsoid in shape and between 0.2 to 0.5 um Feret diameter (largest axis length). Bleached rounded and smaller (<0.1 um) organelles within the melanophores corresponded to typical pre-melanosomes still in process of differentiation.

5.3.2.2 Background colouration

Two chromatophores types were found in light-brown skin portions (Fig. 5-1c). Melanophores were found deeper and closer to the dermal area while xanthophores where at the superficial level, underneath the epidermis. The cv present within xanthophores were different in appearance to those in coloured skin: they were densely compacted, much more uniform in shape and size (0.76 – 1.1 um) and there were no visible pt within xantophores. The cytoplasmic characteristics of melanophores from coloured skin (Fig. 5-1a and 5-1b) where not

different from light-brown skin (Fig. 5-1.c) and their position remained beneath the xantophores. Black-background skin tissue revealed the presence of only melanophore-like structures. Several melanophores (or part of them) are arranged in parallel with an abundant number of ms within the melanophore mainly ellipsoid in shape (Fig. 5-1d) with sizes ranging from ~0.1-0.4 um.

5.3.2.3 Morphologic variation in dermal melanosomes

I analysed individual ms (n=970) from between 7-10 cuts per type of tissue (dark-black and light-brown backgrounds, plus their respective coloured spot; Fig. 5-1). Several patterns were observed at the ms morphology level between different skin tissues. First, the highest level of ms aggregation (density) was found in coloured spots skin portions (>61%) and no statistically difference was found between frog types (dark- vs- light brown, $p>0.5$) (Fig. S5A.2, Table 5-2). Conversely, density level was significantly lower in dark-black than in light-brown skin portions ($\bar{x}$=59% ± 0.004, Fig. S5A.2a, $p<0.01$) indicating that ms in the dark phenotype are somehow less aggregated within melanophores. Second, there was a consistent and significant increase in ms size (~1 fold in area) in skin portions of the dark phenotype ($\bar{x}$=0.25± 0.005 um^2) when compared with the light-brown one ($\bar{x}$=0.12± 0.003 um^2, Fig. S5A.2b). Area of ms in coloured spots portions was similar to that of the light-brown phenotype (~0.14 um^2) and there was no statistically difference between ms area of coloured spots of the dark and light phenotypes (Fig. S5A.2b). Finally, ms were significantly ($p<0.05$) more ellipsoid in shape (circularity value of 1 indicates a perfectly rounded particle) in the dark-black phenotype (~0.78) than in the light-brown one (~0.84) regardless of the evaluated skin portions (background vs. coloured spots)(Fig. S5A.2c, Table 5-2).

Table 5-2. Summary statistics (mean and standard error) of three melanosome-related variables (n=970). Mean values with the same letter are not statistically different (T-test, P>0.05)

Type of tissue	Density (%)		ms area (um^2)		Circularity	
	$\bar{x}$	s.e	$\bar{x}$	s.e	$\bar{x}$	s.e
Dark-black (background)	0.596 - A	0.0041	0.246 - A	0.0046	0.780 - A	0.0048
Dark-black (coloured spots)	0.615 - C	0.0045	0.142 - C	0.0093	0.74 - A	0.0102
Light-brown (background)	0.606 - B	0.0035	0.119 - B	0.0030	0.827 - B	0.0054
Light-brown (coloured spots)	0.617 - C	0.0036	0.149 - C	0.0047	0.848 - B	0.0087

Table 5-3. Estimated genetic distance over MC1R sequence pairs between species. Standard error estimates are shown above the diagonal (bootstrap method, n=999). Analyses were conducted using the Tamura 3-parameter model as suggested by jModelTest (see methods). Bold values indicates the min. and max. genetic distances.

	Lineage I	Lineage II	Lineage III	*O. lehmanni*	Lineage IV
Lineage I		0.0006	0.0010	0.0046	0.0043
Lineage II	0.0016		0.0010	0.0045	0.0042
Lineage III	0.0022	0.0027		0.0046	0.0043
O. lehmanni	0.0183	0.0183	**0.0191**		0.0006
Lineage IV	0.0176	0.0175	0.0184	**0.0014**	

5.3.3 Sequence and evolutionary analyses of MC1R

I obtained Sanger sequences for a total of 64 individuals. Based on the predicted protein homology with *R. temporaria*, I identified the predicted the N-terminus, the internal/external loops, the 7 transmembrane domains, and the C-terminus of the MC1R gene. Within the *O. histrionica* complex, the corrected genetic distances (substitution model=T92+I) ranged from 0.0014 (Linage III - *O. lehmanni)* to 0.019 (Lineage II – *O. lehmanni*) (Table 5-3). After allele phasing, I found a total of 19 different alleles defined by 26 segregating sites ($S_{\text{synonymous}}$= 11, $S_{\text{non-synonymous}}$= 15) in the *O. histrionica* complex (Table 5-4).

A total of 8 alleles were found in the *Linage II*, characterized by light-brown background colouration. Among them, 6 alleles encode a predicted full length (FL, 318 aa, 7-

transmembrane domains) MC1R protein, while the alleles *Hap-3* and *Hap-15* showed an insertion at position 433-bp (Δ433) predicted to cause a frame-shift leading to a truncated MC1R receptor (150 aa, 3-transmembrane domains. Fig. 5-4, Table 5-4). Every genotyped individual from *Linage II* (n= 22) carried at least one copy of a predicted full length (FL) allele. Although the majority of the frogs carried two copies of FL alleles, nine individuals were FL/Δ433 heterozygous. The background colouration of these heterozygous individuals varied across populations. While in Pacurita and Tutunendo FL/Δ433 frogs showed light-brown background colouration similar to that of full length homozygous, in Angostura FL/Δ433 individuals showed a black/light brown variegated pattern (Figure S5A.3). Δ433 alleles appeared to explain the black background colouration that characterizes Linage I specimens as all individuals carried two Δ433 copies. Further support for the possible involvement MC1R alleles in the colouration of Harlequin poison frogs comes from the genotyping results of *Linage IV* and *O. lehmanii*. In these two black southern lineages all individuals carried two copies of predicted truncated alleles (150 aa and 3-transmembrane domains), in this case resulting from a transition (C>A) at the position 432-bp (C432A; Fig. 5-4, Table 5-4).

All genotyped *Linage II* individuals were either homozygous for Δ433 or heterozygous for Δ433 /full-length. Although variation in colour from dark brown to black was not apparent, the multivariate analysis of colouration showed an association between the presence/absence of full-length alleles and the colouration of the dorsal skin (P=0.048) (Fig. 5-4). The two individuals of the outgroup *O. sylvatica*, showing a light brown-reddish dorsal colouration were both homozygous for a single allele encoding a predicted full-length (318 aa) MC1R protein (Fig. 5-2).

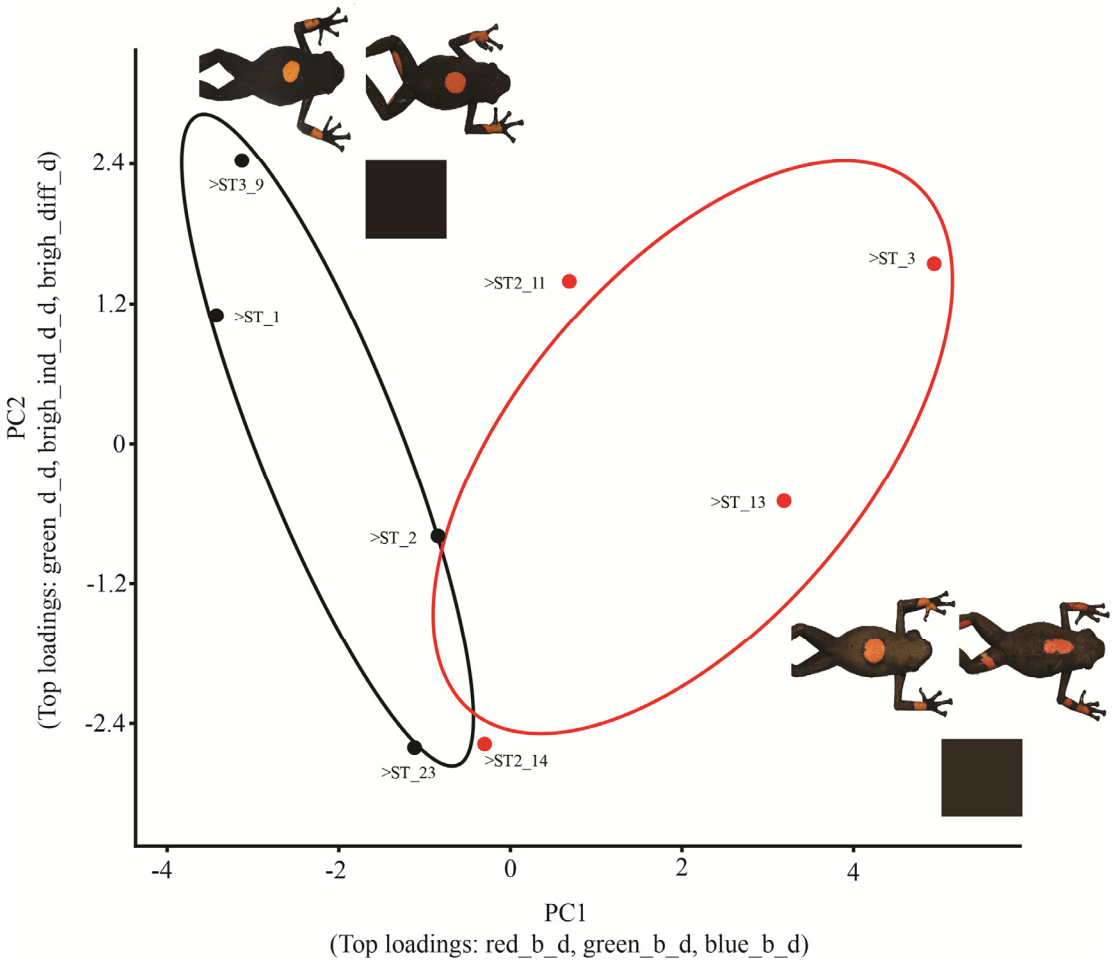


Figure 5-4. PCA analysis of eight Linage II genotyped individuals based on 12 colour variables. Homozygous individuals showing the Δ433 mutation (left) were statistically different (P=0.048) to those where a full-length allele was present (right). Solid coloured boxes represent the colouration intensity of the average RGB values on each group of individuals.

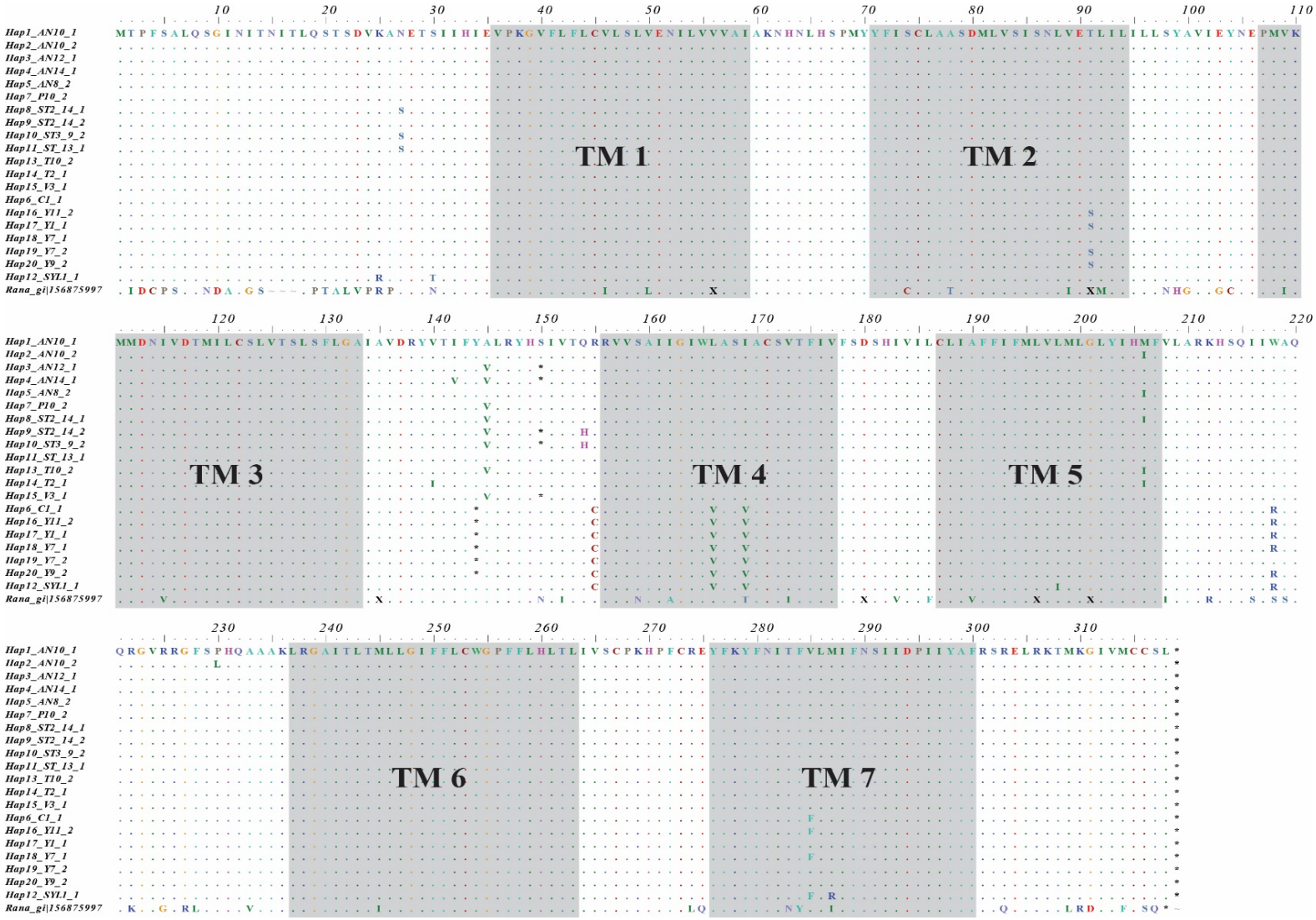


Figure 5-5. Aminoacid alignment of the different MC1R alleles found in this study. Asterisk represent stop codons and are found at positions 150 (Lineages I, II and III*)* and 144 (Lineage IV and *O lehmanni*). MC1R Aminoacid sequence of *Rana catesbiana* was included as a reference. TM= transmembrane domains.

Table 5-4. Haplotype description and nucleotide mutations found at 19 different haplotypes in *Oophaga* lineages (excluding *O. sylvatica*). Shadowed cells indicate haplotypes associated with dark-black background colouration (Fig. 5-1). n/s= non-synonymous, s=synonymous, n-s=nonsense mutations.

Hap.	Population	Lineage	Nucleotide Position																								
			80	**180**	**270**	**271**	**387**	**390**	**418**	**424**	**432**	**433***	**435**	**463**	**464**	**497**	**506**	**550**	**607**	**619**	**622**	**653**	**690**	**766**	**769**	**781**	**787**	**854**
Hap_1	Barranquito, Santa Cecilia, Angostura	Linage II, Linage III	A	G	G	A	A	T	G	A	**C**	**-**	C	G	C	T	A	T	C	G	C	T	C	T	T	C	T	G
Hap_5	Barranquito, Angostura	Linage II	.	.	.	.	.	.	.	.	.	**-**	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
Hap_2	Angostura	Linage II	.	.	.	.	.	.	.	.	.	**-**	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.
Hap_7	Pacurita	Linage II	.	.	.	.	.	.	.	.	.	**-**	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_13	Tutunendo	Linage II	.	.	.	.	.	.	.	.	.	**-**	T	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
Hap_14	Tutunendo	Linage II	.	.	.	.	.	.	A	.	.	**-**	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.
Hap_8	Santa Cecilia	Linage III	G	.	.	.	.	.	.	.	.	**-**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_3	Bahia, Victoria, Salero, Pacurita, Tutunendo, Angostura	Linage I, Linage II	.	.	.	.	.	.	.	.	.	**A**	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_15	Valle, Victoria	Linage I	.	A	.	.	.	.	.	.	.	**A**	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_4	Angostura	Linage II	.	.	.	.	.	.	.	G	.	**A**	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_9	Santa Cecilia	Linage III	G	.	.	.	.	.	.	.	.	**A**	T	C	.	.	.	.	.	A	.	.	.	.	.	.	.	.
Hap_10	Santa Cecilia	Linage III	.	.	.	.	.	.	.	.	.	**A**	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_11	Santa Cecilia	Linage III	G	.	.	.	.	.	.	.	.	**A**	T	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_6	*O. lehmanni,* Cauchal	*O. lehmanni,* Linage IV	.	.	A	.	T	C	.	.	**A**	**-**	.	.	T	G	G	C	G	.	T	A	.	A	C	T	C	T
Hap_16	Anchicaya	Linage IV	.	.	A	T	T	C	.	.	**A**	**-**	.	.	.	G	G	C	G	.	T	A	.	A	C	T	C	T
Hap_17	Anchicaya	Linage IV	.	.	A	T	T	C	.	.	**A**	**-**	.	.	.	G	G	C	.	.	T	A	.	A	.	.	.	.
Hap_18	Anchicaya	Linage IV	.	.	A	.	T	C	.	.	**A**	**-**	.	.	.	G	G	C	G	.	T	A	.	A	C	T	.	T
Hap_19	Anchicaya	Linage IV	.	.	A	T	T	C	.	.	**A**	**-**	.	.	.	G	G	C	G	.	.	.	.	.	.	.	.	.
Hap_20	Anchicaya	Linage IV	.	.	A	T	T	C	.	.	**A**	**-**	.	.	.	G	G	C	G	.	T	A	.	A	C	.	C	.
	Substitution type		n/s	s	s	n/s	s	s	n/s	n/s	**n-s**	**n-s**	n/s	n/s	n/s	n/s	n/s	s	s	n/s	s	n/s	n/s	s	s	s	s	n/s

The evolutionary relationships between MC1R alleles (ML, BI, and haplotype network) is shown in Figure 5-2. All analyses were congruent and revealed three main clades of alleles associated with both, geographic location and background colouration. The southern clade of alleles included all the C432A alleles found in Linage IV and *O. lehmanni*. These alleles encoding 3-transmembrane domain receptors are more related to the full length alleles found in *O. sylvatica* than to the Δ433 alleles encoding 3-transmembrane receptors in the northern range of complex. Δ433 are very closely related to the FL alelles and other than the insertion at 433-bp there are no fixed differences between them.

Tajima's D were not significant (P>0.10) for the whole set of sequences (D=1.49) neither when analysed by independent lineages (Lineage II= 0.99; Lineage I= 1.19; Lineage III= 1.26; Lineage IV-*lehmanni*=-0.4). Fu and Li's tests showed a similar pattern when analysed by species, however, this test was significant when evaluated overall sequences (D*=2.05, P<0.02). Accordingly, overall Z-tests of selection showed no significant deviations from neutrality (Table S5B.2). Although the MK-tests revealed and excess of synonymous polymorphisms none of them were statistically significant.

5.4 Discussion

5.4.1 Structural basis of colour variation

Three types chromatophores are known to be related with both permanent and temporary colouration in amphibians: melanophores, containing black melanin; xanthophores, containing yellow/orange coloured pigments; and iridophores, containing reflective or iridescent platelets (Bagnara & Hadley, 1969; Bagnara *et al.*, 1979). In the few Anuran species studied thus far, three types of cells are usually organized in the so called dermal chromatophore unit such that

the xanthophores are located on uppermost layer of the dermis, iridophores are situated under that, and the melanophores are located at the very bottom (Bagnara *et al.*, 1968) . In contrast, the microscopy analyses revealed that the dorsal skin of *Oophaga* frogs does not necessarily contain all these structural components. Specifically, no evidence for the presence of iridophores was found. A reduced number of iridiophores imparts a bright red and/or yellow colouration (Frost & Robinson, 1984) which is a common characteristic of the *Oophaga* species studied here. Thus, it is very likely that the skin of harlequin poison frogs simply lacks these type of chromatophores. However, although less likely, it is also possible that a few iridiophores were present but that they were shattered when sectioning the embedded tissue (see Bagnara *et al.*, 1968 for a detailed explanation).

The histological analysis presented here clearly document the structural basis of the dorsal colour differences between background and coloured-spots in *Oophaga* poison frogs. The dorsal spots can be explained by the presence of an outer layer of xanthophores with conspicuous pterinosomes and relatively small carotenoid-containing vesicles that are located over a single layer of melanophores with densely aggregated melanosomes probably towards the nucleus of the melanophore (Frost & Robinson, 1984). The structural composition and subcellular characteristics of the coloured spots appeared to be identical between individuals showing either black or light-brown background colouration.

Dorsal light-brown background colouration is explained by the presence of xantophores and melanophores organized in a similar pattern to that of the coloured spots. However, there were several melanophores layers located as parallel stacks underneath the xantophores and the characteristics of the organelles are quite different in both type of cells: in the melanophores, the melanosomes were slightly less packed around the nucleus while in the xantophores, there

were no pterinosomes and the carotenoid-containing vesicles were larger which may cause dilution effect of carotenoid pigments. It is well known that aggregated melanosomes are associated with lighter phenotypes in at least one frog species (Snezhko *et al.*, 2010) and that long-term adaptation of fishes to lighter backgrounds also involves melanosome aggregation (Sugimoto, 2002). Thus, the lack of red pteridines coupled with more diluted brownish-yellow carotenes and relatively packed melanosomes are likely to explain the pale-brown phenotype of *Linage II* specimens . Taken all together my results reveal that dark-black dorsal colouration is explained by a combination of different cytological and histological factors. First, black skin is exclusively composed of melanophore layers. Second, the melanosome size is ~1 fold bigger than in any other type of skin section and third, the ms are sparsely and uniformly distributed within the melanophores.

5.4.2 Background colour variation and MC1R mutations

After analysing and comparing the MC1R DNA sequences in 64 individuals of several *Oophaga* species, the most interesting mutations were those that were associated with truncated alleles (Δ433, C432A) whose predicted products lack most of the domains including several internal- external loops (IL2, IL3, EL2, EL3), transmembrane regions (TM4-7) and the intracellular tail (Klovins *et al.*, 2004). Due the lack of these important functional components, it is reasonable to consider that the gene products of these alleles may not be functional. In other species, the presence of frameshifts, deletions or early stop codons affecting important portions of the MC1R protein are suggested to inactivate the receptor function causing a continuous production of different melanin pigments and hence, different resulting contrasting phenotypes (Robbins *et al.*, 1993; Klungland *et al.*, 1995; Klungland *et al.*, 1999; Våge *et al.*, 1999; Everts *et al.*, 2000; Newton *et al.*, 2000; Fontanesi *et al.*, 2009; Gross *et al.*, 2009). In these homeotherms

organisms, the lack of function at the MC1R level has been associated with colourations from albino to reddish spots across the body, including pale-brown and reddish hair colouration. Although the vast majority of research involving colour variation and the effects of MC1R in colour change has been performed in a mammalian and avian context, there is no apparent reason to believe that a complete different mechanism is involved in poikilothermic vertebrates (fish and amphibians). In contrast to mammals and birds, poikilothermic vertebrates have multiple pigment cells (chromatophores, Bagnara *et al*, 1986) that rather than transferring pigments to other cell types, they retain their pigments intracellularly. Despite these differences, many of the genetic pathways critical for melanocyte differentiation and morphogenesis are conserved in melanophores and other poikilothermic chromatophore types (Mills & Patterson, 2009). The findings of this study in *Oophaga* poison frogs are clearly contrasting with those previously described in mammals and birds. Here, I found a perfect association between melanised dark-black background phenotypes and highly truncated MC1R alleles, and conversely, complete 7-transmembrane domains MC1R receptors were associated with light-browned background individuals.

It is well known that many amphibians undergo dramatic colouration and patterning changes during metamorphosis due hormonal changes. Dark hues are known to be produced by the interaction between high levels of melanocyte-stimulating hormone (α-MSH) and several variants of its transmembrane receptor (MC1R) through the dispersion of melanosomes within the melanophore (by increasing cAMP intracellular levels) (Sugimoto, 2002; Logan *et al.*, 2006). In this study, the perfect association between highly truncated (and likely not functional) MC1R receptors and dark-black phenotypes may be an indicator of other genes playing a role in poikilothermic dark hues (e.g melanin-concentrating hormone MCH) (Sugimoto, 2002; Logan *et*

al., 2006; Gross *et al.*, 2009). However, the presence of certain levels of cellular function in highly truncated G-receptors cannot be discarded. It is known that G-coupled proteins (GPCR) as MC1R, may increase the total number of receptor isoforms by alternative splicing (Wise, 2012). Many GPCRs that are retained in the endoplasmic reticulum can still be functionally active, as well, highly truncated receptors may conserve their function despite the loss of important functional components that generates as low as only two transmembrane domains (Xie *et al.*, 2000; Córdoba-Chacón *et al.*, 2010; Wise, 2012). In any case, this study did not aimed to disentangle the physiological and biochemical consequences of truncated MC1R variants. Studies at the gene expression and functional analysis levels must be performed in order to understand the implications in colour variation of the highly truncated MC1R receptors found in this study.

5.4.3 Genetics basis of background colouration

My results also provide initial insights about the genetic underpinnings of colouration in these frogs. These findings suggest a predominant monogenic-dominant effect of MC1R alleles of light-brown colouration in *Oophaga* poison frogs. More than 50% of the individuals with light-brown phenotype were homozygous for 7-domains MC1R versions while heterozygous individuals with the Δ433 mutation retained lighter phenotypes (Fig. 5-6). Accordingly, homozygous and heterozygous individuals (Fig. S4b) with alleles showing either, the Δ433 and C432A mutations (both truncated MC1R receptors) were associated with the dark-black background colouration, more dispersed melanosomes and increased melanosome size. Based on these results, it seems reasonable to consider that the presence of one copy of a complete 7-domains MC1R allele is enough to trigger lighter phenotypes while truncated alleles in homozygous state produce the dark-black ones (Fig. 5-4). Similar results have been found in *O.*

pumilio, a closely related species. By using molecular markers to estimate pedigree relationships among individuals from a polymorphic wild population, it was demonstrated that variation in dorsal colouration could be explained by a single locus with dominance (Richards-Zawacki *et al.*, 2012). I emphasize that the analysis of variation in natural populations *per se* must be taken with caution and hence, this hypothesis remains speculative until further experimental crosses between contrasting phenotypes are performed.

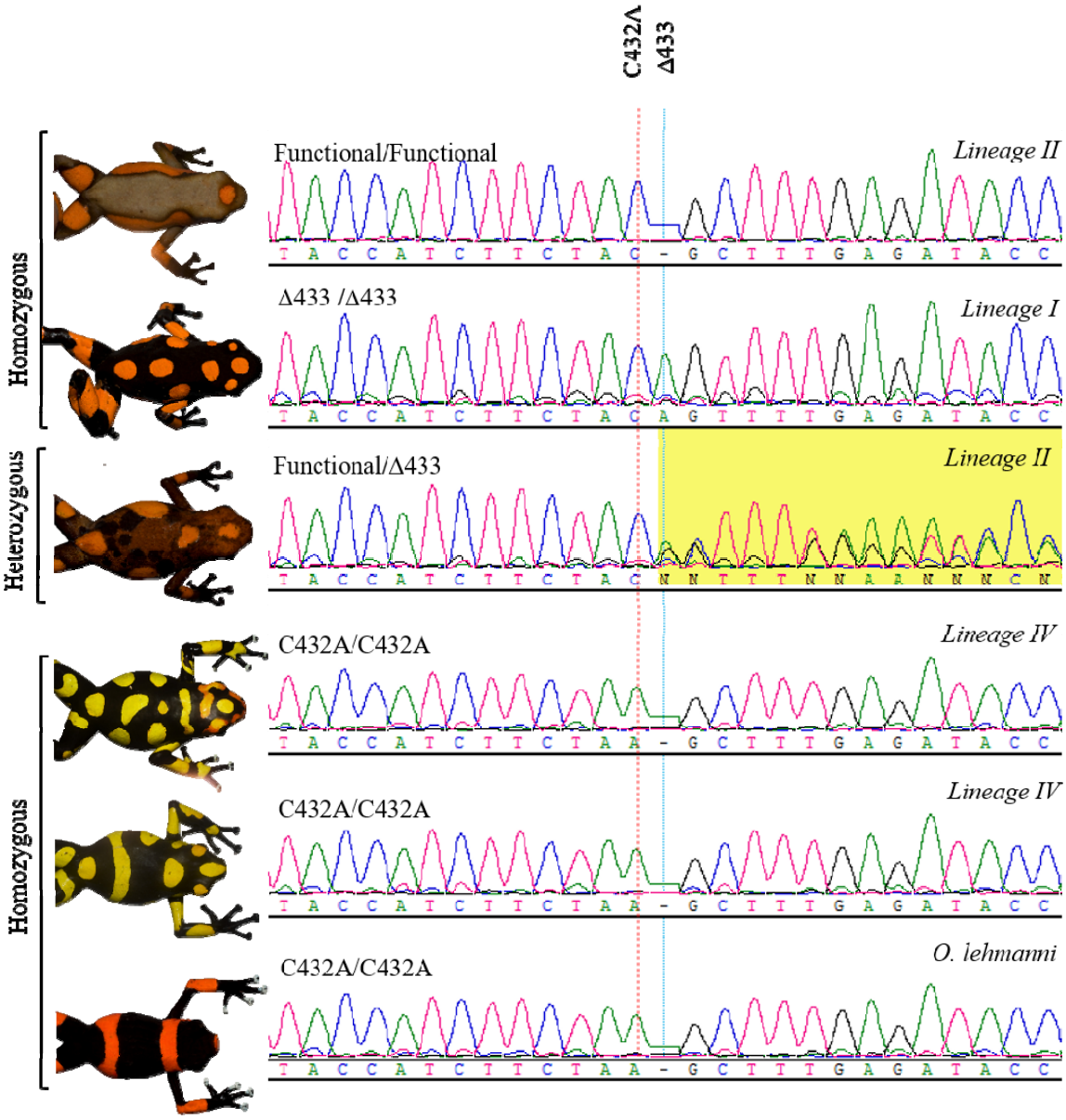


Figure 5-6. Representation of MC1R allele variation and its association with background colouration. Mutation C432Ais represented by a red-dotted line while mutation Δ433 is indicated in blue. Both mutations generate highly truncates MC1R receptors. (144 and 150 aa). We have designated as "functional" any allele without early stops codons producing 7-domains MC1R proteins (318 aa). Heterozygous individuals with only one copy of the complete MC1R protein retain the light-brown colouration in most of the cases.

One of the most interesting findings of this study is the observation of similar phenotypes appearing independently in nature (convergent evolution). I found that the dark-black background phenotype has arisen independently in geographically separate lineages (Lineage I

vs. Lineage IV – *O. lehmanni*) through different mutations of the same gene that, in both cases, generate truncated MC1R receptors (Fig. 5-5, Table 5-4). The convergence through different mutations with similar effects may be more likely towards a complete or partial loss of function. That is, a pathway or gene function could be disrupted with more probability by mutations, as opposed to a gain-of-function. In aposematic organisms, warning signals are proposed to be detectable and easily learnt by predators if they show a strong colour contrast against the background on which they occur (Prudic *et al.*, 2007). The chromatic-contrast analysis of this study showed that the darker background phenotypes ($\bar{x}$=89.57± 1.86 and $\bar{x}$=108.90± 2.48) may probably serve as a more detectable signal when compared to the lighter background ($\bar{x}$=8.14± 4.99) due the markedly brightness contrast between the coloured spots and the background colouration (Amézquita *et al.*, 2013; Medina *et al.*, 2013; Galindo-Uribe *et al.*, 2014). Even though direct evidence of selection patterns was not found in this study (dark-truncated vs. light-complete MC1R variants), it is still possible to considered that the dark-black background phenotypes may confer certain levels of fitness advantages on these species, probably through an increased and more efficient aposematic signal (Gamberale-Stille & Guilford, 2003; Prudic *et al.*, 2007; Aronsson & Gamberale-Stille, 2009, 2012). Hence, the aposematic advantage could be responsible for the appearance of the dark-black phenotype independently in different geographically isolated evolutionary lineages. This example of convergent evolution in adaptive traits has been reported in a wide variety of organisms (Summers, 2003; Vences *et al.*, 2003; Hoekstra *et al.*, 2006; Gross *et al.*, 2009; Steiner *et al.*, 2009; Rosenblum *et al.*, 2010; Kaeuffer *et al.*, 2012; Ellingson *et al.*, 2014) and together with the evidence provided in this study, it indicates that MC1R has been a frequent target of mutation in the evolution of contrasting colour phenotypes in amphibians and fishes (Gross *et al.*, 2009; Mundy, 2009).

Finally, this study adds to a growing body of literature showing that it is surprisingly common that similar phenotypes evolving in parallel share the same genetic mechanisms and constitutes a documented example of genetic repeatability. Overall, this study provides evidence to suggest that variation at the MC1R receptor could potentially be a major factor responsible for the high phenotypic variation of aposematic signals of Harlequin poison frogs. As mentioned above, further experimental genetic crosses, QTL mapping analysis of chromatophore characteristics and molecular functional analysis are required to confirm the phenotypic implications of MC1R variants in the colouration of these highly endangered frogs.

CHAPTER 6

GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

In general, poison frogs of the *Oophaga* genus, including the members of the Histrionica complex are still an intriguing and understudied system in the field of herpetology, ecology and evolutionary biology. The vast majority of literature in this taxa has focused in only one Panamanian species (*O. pumilio*) (Daly *et al.*, 2003; Hagemann & Pröhl, 2007; Rudh *et al.*, 2007; Hauswaldt *et al.*, 2009; Brown *et al.*, 2010; Hegna *et al.*, 2011; Richards-Zawacki *et al.*, 2012), probably because of the striking phenotypic variation in colouration and patterning of populations from the "Bocas del Toro" archipelago. Only a few attempts have been made to study other species of the genus, including *O. granulifera* (Brusa, 2012; Brusa *et al.*, 2013; Willink *et al.*, 2013) and *O. histrionic*a (Amézquita *et al.*, 2013; Medina *et al.*, 2013; Vargas-Salinas & Amézquita, 2013). By using a multi-disciplinary approach that combined phenotypic, ecological and molecular genetic analysis, this PhD dissertation and its further outcomes constitute the first comprehensive study of the Histrionica complex. The main and general contributions of this study are presented below, in the same order as they appeared throughout the main document (Chapters 2-5).

The findings presented in Chapter 2 indicated that *Oophaga* frogs originated in Central America and reached South America after the closure of the Panama Isthmus. The South (*O. sylvatica, O. histrionica, O. lehmanni*) and Central (*O. granulifera, O. pumilio, O. speciosa, O. arborea, O. vicentei*) American clades of this genus have convergently evolved to similar patterns of geographic distribution and niche occupancy. Within the Central and South American clades, sister taxa showed parapatric distributions replacing each other along elevation gradients, as predicted by the models of divergence along continuous ecological gradients. Accordingly, I

found strong shifts in climatic niches throughout the history of these two clades. However, the largest niche shifts seem to post-date the final elevation of the Talamanca and northern Andes. Overall, these data strongly suggest that speciation along climatic gradients on a structured landscape has been a major evolutionary force behind the diversification of *Oophag*a poison frogs.

Early taxonomic studies (Berthold, 1846; Myers & Daly, 1976) and genetic data from geographically scattered individuals (Grant *et al.*, 2006; Brown *et al.*, 2011) have failed to recognized the striking phenotypic diversity of Harlequin poison frogs. In Chapter 3, by applying an iterative protocol for lineage delimitation under the generalized lineage concept (GLC) (de Queiroz, 1998, 2007), I showed that the resulting groups should correspond to evolutionarily independent metapopulation lineages because they reflect the common signal of different secondary defining properties (environmental and genetic distinctiveness, phenotypic diagnosability, etc.), implying the existence of barriers preventing or limiting gene exchange. This analyses revealed the existence of at least five different lineages within the Histrionica complex, some of them occurring in very small isolated populations outside any protected areas. More broadly, this study exemplifies how transcriptome-based reduction of genomic complexity and multivariate statistical analysis of environmental and phenotypic traits can be integrated to successfully identify independent evolutionary lineages and their boundaries. These results have important conservation implications as some of its members are considered amongst the most endangered species of all amphibians. Regardless of their conservation status, none of the *Oophaga* lineages described in Chapter 3 occur in protected areas and ideally, once included in the appropriate endangered list, these charismatic frogs will serve as umbrella species for the future conservation of entire ecosystems.

In Chapter 4, transcriptomics analyses offered an ideal opportunity to further our understanding of the alkaloid sequestration system, auto-resistance to toxic compounds and variation in colouration and pattern in Harlequin poison frogs. To date, no studies addressing the molecular basis or genetic architecture of these important mechanisms had been adopted. By taking advantage of the high coverage offered by NGS, this study allowed me to (i) use RNA-sequencing to generate an annotate a *de-novo* skin transcriptome of Harlequin poison frogs; (ii) conduct a comparative transcriptome analysis of four *Oophaga* lineages; (iii) identify candidate genes for important adaptive traits and (iv) construct a genomic phylogeny for these polymorphic species. The findings of this Chapter provided a valuable genomic resource for the amphibians research and contribute insights into the molecular mechanisms of important adaptive traits in *Oophaga* frogs. This study reports the first comprehensive study of *Oophaga* skin transcriptomes and pave the road for future diversity and functional studies in this charismatic group of species.

In Chapter 5, I studied the ultracellular structural basis of colour and the role of a candidate gene in colour variation. I showed that an important component of aposematic signals (background colouration) is explained at the subcellular level by the absence of a "typical" dermal chromatophore unit (Bagnara *et al.*, 1968). More than 30 years after the pioneer works describing the histological basis of colouration in amphibians, this study represent the first attempt for unravelling the structural basis of colouration in aposematic frogs. Without any doubt, further histological studies including a wider spectrum of colour and pattern variation should be performed. This field of biology offers exciting possibilities and new studies could provide the initial clues towards unravelling the molecular and genetics basis of aposematic and cryptic signalling in amphibians. Finally, by using a candidate-gene approach, I provided

preliminary hypothesis of the functional effect of MC1R mutations and the genetic mechanisms of colour variation in this frogs. Furthermore, one of the most interesting findings of this study is the observation of similar phenotypes appearing independently in nature as a consequence of different mutations of the same gene, a classical example of convergent phenotypic evolution and genetic repeatability. Broadly, this part of my dissertation provided evidence to suggest that variation at the MC1R receptor could potentially be a major factor responsible for the high phenotypic variation of aposematic signals of Harlequin poison frogs. Though this work provided initial insights about the genetic and molecular basis of background colour variation, further experimental genetic crosses, QTL mapping analysis of chromatophore characteristics and molecular functional analysis are required to confirm the phenotypic implications of MC1R variants.

LIST OF REFERENCES

Abratis, M. & Wörner, G. (2001) Ridge collision, slab-window formation, and the flux of Pacific asthenosphere into the Caribbean realm. *Geology*, **29**, 127-130.

Achard, F., Eva, H.D., Stibig, H.J., Mayaux, P., Gallego, J., Richards, T. & Malingreau, J.P. (2002) Determination of deforestation rates of the world's humid tropical forests. *Science*, **297**, 999-1002.

Alaluf, S., Atkins, D., Barrett, K., Blount, M., Carter, N. & Heath, A. (2002) Ethnic Variation in Melanin Content and Composition in Photoexposed and Photoprotected Human Skin. *Pigment Cell Research*, **15**, 112-118.

Alaluf, S., Heath, A., Carter, N.I.K., Atkins, D., Mahalingam, H., Barrett, K., Kolb, R.I.A. & Smit, N. (2001) Variation in Melanin Content and Composition in Type V and VI Photoexposed and Photoprotected Human Skin: The Dominant Role of DHI. *Pigment Cell Research*, **14**, 337-347.

Alho, J., Herczeg, G., Soderman, F., Laurila, A., Jonsson, K. & Merila, J. (2010) Increasing melanism along a latitudinal gradient in a widespread amphibian: local adaptation, ontogenic or environmental plasticity? *BMC Evolutionary Biology*, **10**, 317.

Allison, E.G. & Funk, W.C. (2009) Sexual Selection on Morphology in an Explosive Breeding Amphibian, the Columbia Spotted Frog (Rana luteiventris). *Journal of Herpetology*, **43**, 244-251.

Amézquita, A., Flechas, S.V., Lima, A.P., Gasser, H. & Hödl, W. (2011) Acoustic interference and recognition space within a complex assemblage of dendrobatid frogs. *Proceedings of the National Academy of Sciences of the United States of America Biological Sciences*,

Amézquita, A., Castro, L., Arias, M., González, M. & Esquivel, C. (2013) Field but not lab paradigms support generalisation by predators of aposematic polymorphic prey: the *Oophaga histrionica* complex. *Evolutionary Ecology*, **27**, 769-782.

Anacker, B.L. & Strauss, S.Y. (2014) The geography and ecology of plant speciation: range overlap and niche divergence in sister species. *Proceedings of the Royal Society Biological Sciences Series B*, **281**, 20132980.

Andrés, J.A. & Bogdanowicz, S.M. (2011) Isolating Microsatellite Loci: Looking Back, Looking Ahead. *Molecular Methods for Evolutionary Genetics* (ed. by V. Orgogozo and V.M. Rockman), pp. 211-232. Humana Press, Totowa, NJ.

Aronsson, M. & Gamberale-Stille, G. (2009) Importance of internal pattern contrast and contrast against the background in aposematic signals. *Behavioral Ecology*, **20**, 1356-1362.

Aronsson, M. & Gamberale-Stille, G. (2012) Evidence of signaling benefits to contrasting internal color boundaries in warning coloration. *Behavioral Ecology*,

Auber, L. (1957) The distribution of structural colours and unusual pigments in the class Aves. *Ibis*, **99**, 463-476.

Bagnara, J. & Taylor, J. (1970) Differences in pigment-containing organelles between color forms of the red-backed salamander, Plethodon cinereus. *Zeitschrift für Zellforschung und Mikroskopische Anatomie*, **106**, 412-417.

Bagnara, J.T. (1982) Development of the spot pattern in the leopard frog. *Journal of Experimental Zoology*, **224**, 283-287.

Bagnara, J.T. (2003) Enigmas of Pterorhodin, a Red Melanosomal Pigment of Tree Frogs. *Pigment Cell Research*, **16**, 510-516.

Bagnara, J.T. & Hadley, M.E. (1969) The control of bright colored pigment cells of fishes and amphibians. *American Zoologist*, **9**, 465-478.
Bagnara, J.T., Taylor, J.D. & Hadley, M.E. (1968) The dermal chromatophore unit. *The Journal of Cell Biology*, **38**, 67-79.
Bagnara, J.T., Taylor, J.D. & Prota, G. (1973) Color Changes, Unusual Melanosomes, and a New Pigment from Leaf Frogs. *Science*, **182**, 1034-1035.
Bagnara, J.T., Frost, S.K. & Matsumoto, J. (1978) On the Development of Pigment Patterns in Amphibians. *American Zoologist*, **18**, 301-312.
Bagnara, J.T., Matsumoto, J., Ferris, W., Frost, S.K., Turner, W.A., Jr., Tchen, T.T. & Taylor, J.D. (1979) Common Origin of Pigment Cells. *Science*, **203**, 410-415.
Baião, P.C. & Parker, P.G. (2012) Evolution of the Melanocortin-1 Receptor (MC1R) in Boobies and Gannets (Aves, Suliformes). *Journal of Heredity*, **103**, 322-329.
Barley, A.J., White, J., Diesmos, A.C. & Brown, R.M. (2013) The challenge of species delimitation at the extremes: diversification without morphological change in Philippine sun skinks. *Evolution*, **67**, 3556-3572.
Bauer, L. (1994) New names in the family Dendrobatidae (Anura, Amphibia). *RIPA Fall*, **1994**, 1-6.
Baum, D.A. & Donoghue, M.J. (1995) Choosing among Alternative "Phylogenetic" Species Concepts. *Systematic Botany*, **20**, 560-573.
Becerro, M., Starmer, J. & Paul, V. (2006) Chemical Defenses of Cryptic and Aposematic Gastropterid Molluscs Feeding on their Host Sponge Dysidea granulosa. *Journal of Chemical Ecology*, **32**, 1491-1500.
Berlocher, S.H. & Feder, J.L. (2002) Sympatric speciation in phytophagous insects: Moving Beyond Controversy? *Annual Review of Entomology*, **47**, 773-815.
Berthold, A.A. (1846) Ueber verschiedene neue oder seltene Reptilien aus Neu-Granada und Crustaceen aus China. *Nachr. Ges. Wiss. Göttingen*, **3**, 43.
Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, **22**, 148-155.
Bolivar, W., Castro, F. & Lötters, S. (2004) *Oophaga lehmanni*. The IUCN Red List of Threatened Species. *e.T55190A11255085.*,
Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, **10**, e1003537.
Bowdish, T.I. & Bultman, T.L. (1993) Visual cues used by mantids in learning aversion to aposematically colored prey. *American Midland Naturalist*, 215-222.
Brown, D., Brenneman, R., Koepfli, K.-P., Pollinger, J., Mila, B., Georgiadis, N., Louis, E., Grether, G., Jacobs, D. & Wayne, R. (2007) Extensive population genetic structure in the giraffe. *BMC Biology*, **5**, 57-70.
Brown, J. (2013) The evolution of parental care, aposematism and color diversity in Neotropical poison frogs. *Evolutionary Ecology*, **27**, 825-829.
Brown, J.L., Maan, M.E., Cummings, M.E. & Summers, K. (2010) Evidence for selection on coloration in a Panamanian poison frog: a coalescent-based approach. *Journal of Biogeography*, **37**, 891-901.
Brown, J.L., Twomey, E., Amézquita, A., Souza, M.B.d., Caldwell, J.p., Lötters, S., May, R.v., Melo-Sampaio, P.R., Mejía-Vargas, D., Perez-Peña, P., Pepper, M., Poelman, E.h.,

Sanchez-Rodriguez, M. & Summers, K. (2011) A taxonomic revision of the Neotropical poison frog genus *Ranitomeya* (Amphibia: Dendrobatidae). *Zootaxa*, **3083**, 1-120.
Brumfield, R.T. & Edwards, S.V. (2007) Evolution into and out of the Andes: a Bayesian analysis of historical diversification in *Thamnophilus antshrikes*. *Evolution*, **61**, 346-367.
Brusa, O. (2012) *Geographical variation in the granular poison frog, Oophaga granulifera: genetics, colouration, advertisement call and morphology*. Dissertation, Tierärztliche Hochschule Hannover, Hannover.
Brusa, O., Bellati, A., Meuche, I., Mundy, N.I. & Pröhl, H. (2012) Divergent evolution in the polymorphic granular poison-dart frog, Oophaga granulifera: genetics, coloration, advertisement calls and morphology. *Journal of Biogeography*,
Brusa, O., Bellati, A., Meuche, I., Mundy, N.I. & Pröhl, H. (2013) Divergent evolution in the polymorphic granular poison-dart frog, *Oophaga granulifera*: genetics, coloration, advertisement calls and morphology. *Journal of Biogeography*, **40**, 394-408.
Cardoso, A., Serrano, A. & Vogler, A.P. (2009) Morphological and molecular variation in tiger beetles of the Cicindela hybrida complex: is an 'integrative taxonomy' possible? *Molecular Ecology*, **18**, 648-664.
Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T. & Moritz, C. (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785-789.
Carqueijeiro, I., Noronha, H., Duarte, P., Gerós, H. & Sottomayor, M. (2013) Vacuolar transport of the medicinal alkaloids from Catharanthus roseus is mediated by a proton-driven antiport. *Plant Physiology*, **162**, 1486-1496.
Carstens, B.C., Pelletier, T.A., Reid, N.M. & Satler, J.D. (2013) How to fail at species delimitation. *Molecular Ecology*, **22**, 4369-4383.
Chamala, S., García, N., Godden, G.T., Krishnakumar, V., Jordon-Thaden, I.E., Smet, R.D., Barbazuk, W.B., Soltis, D.E. & Soltis, P.S. (2015) MarkerMiner 1.0: A New Application for Phylogenetic Marker Development Using Angiosperm Transcriptomes. *Applications in Plant Sciences*, **3**, 1400115.
Chaudhary, A. & Willett, K.L. (2006) Inhibition of human cytochrome CYP1 enzymes by flavonoids of St. John's wort. *Toxicology*, **217**, 194-205.
Chikako, M. (2012) Geographic variations of melanocortine 1 receptor gene (MC1R) in the common frog (*Rana temporaria*) in Northern Europe. In: *Amphibia-Reptilia*, pp. 105-111
Chiucchi, J.E. & Gibbs, H.L. (2010) Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. *Molecular Ecology*, **19**, 5345-5358.
Chouteau, M. & Angers, B. (2011) The Role of Predators in Maintaining the Geographic Organization of Aposematic Signals. *The American Naturalist*, **178**, 810-817.
Chouteau, M. & Angers, B. (2012) Wright's Shifting Balance Theory and the Diversification of Aposematic Signals. *PLoS ONE*, **7**, e34028.
Clark, V.C., Raxworthy, C.J., Rakotomalala, V., Sierwald, P. & Fisher, B.L. (2005) Convergent evolution of chemical defense in poison frogs and arthropod prey between Madagascar and the Neotropics. *Proceedings of the National Academy of Sciences*, **102**, 11617-11622.
Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1659.

Clough, M. & Summers, K. (2000) Phylogenetic systematics and biogeography of the poison frogs: evidence from mitochondrial DNA sequences. *Biological Journal of the Linnean Society*, **70**, 515-540.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. & Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674-3676.

Córdoba-Chacón, J., Gahete, M.D., Duran-Prado, M., Pozo-Salas, A.I., Malagón, M.M., Gracia-Navarro, F., Kineman, R.D., Luque, R.M. & Castaño, J.P. (2010) Identification and characterization of new functional truncated variants of somatostatin receptor subtype 5 in rodents. *Cellular and Molecular Life Sciences*, **67**, 1147-1163.

Corso, J., Gonçalves, G. & Thales, d.F. (2012) Sequence variation in the melanocortin-1 receptor (MC1R) pigmentation gene and its role in the cryptic coloration of two South American sand lizards. *Genet. Mol. Biol.*, **35**, 81-87.

Coyne, J.A. & Orr, H.A. (2004) *Speciation*. Sinauer Associates Sunderland, MA.

Crawford, A. (2003) Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution*, **57**, 636-641.

Crothers, L., Gering, E. & Cummings, M. (2011) Aposematic signal variation predicts male-male interactions in a polymorphic poison frog. *Evolution*, **65**, 599-605.

Culver, M., Johnson, W., Pecon-Slattery, J. & O'Brien, S. (2000) Genomic ancestry of the American puma (Puma concolor). *Journal of Heredity*, **91**, 186-197.

Cummings, M. & Crothers, L. (2013) Interacting selection diversifies warning signals in a polytypic frog: an examination with the strawberry poison frog. *Evolutionary Ecology*, **27**, 693-710.

Daly, J.W., Spande, T.F. & Garraffo, H.M. (2005) Alkaloids from Amphibian Skin: A Tabulation of Over Eight-Hundred Compounds. *Journal of Natural Products*, **68**, 1556-1575.

Daly, J.W., Brown, G.B., Mensah-Dwumah, M. & Myers, C.W. (1978) Classification of skin alkaloids from neotropical poison-dart frogs (dendrobatidae). *Toxicon*, **16**, 163-188.

Daly, J.W., McNeal, E.T., Overman, L.E. & Ellison, D.H. (1985) A new class of cardiotonic agents: structure-activity correlations for natural and synthetic analogs of the alkaloid pumiliotoxin B (8-hydroxy-8-methyl-6-alkylidene-1-azabicyclo[4.3.0]nonanes). *Journal of Medicinal Chemistry*, **28**, 482-486.

Daly, J.W., Martin Garraffo, H., Spande, T.F., Jaramillo, C. & Stanley Rand, A. (1994) Dietary source for skin alkaloids of poison frogs (Dendrobatidae)? *Journal of Chemical Ecology*, **20**, 943-955.

Daly, J.W., Ware, N., Saporito, R.A., Spande, T.F. & Garraffo, H.M. (2009) N-Methyldecahydroquinolines: An Unexpected Class of Alkaloids from Amazonian Poison Frogs (Dendrobatidae). *Journal of Natural Products*, **72**, 1110-1114.

Daly, J.W., Kaneko, T., Wilham, J., Garraffo, H.M., Spande, T.F., Espinosa, A. & Donnelly, M.A. (2002) Bioactive alkaloids of frog skin: Combinatorial bioprospecting reveals that pumiliotoxins have an arthropod source. *Proceedings of the National Academy of Sciences*, **99**, 13996-14001.

Daly, J.W., Garraffo, H.M., Spande, T.F., Clark, V.C., Ma, J., Ziffer, H. & Cover, J.F. (2003) Evidence for an enantioselective pumiliotoxin 7-hydroxylase in dendrobatid poison frogs of the genus Dendrobates. *Proceedings of the National Academy of Sciences*, **100**, 11092-11097.

Darst, C.R. & Cummings, M.E. (2006) Predator learning favours mimicry of a less-toxic model in poison frogs. *Nature*, **440**, 208-211.
Darst, C.R., Cummings, M.E. & Cannatella, D.C. (2006) A mechanism for diversity in warning signals: Conspicuousness versus toxicity in poison frogs. *Proceedings of the National Academy of Sciences*, **103**, 5852-5857.
Dayrat, B. (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, **85**, 407-415.
de Queiroz, K. (1998) The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. *Endless forms: species and speciation*, 57-75.
de Queiroz, K. (2005a) Different species problems and their resolution. *BioEssays*, **27**, 1263-1269.
de Queiroz, K. (2005b) Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 6600-6607.
de Queiroz, K. (2007) Species Concepts and Species Delimitation. *Systematic Biology*, **56**, 879-886.
Dejaco, T., Arthofer, W., Sheets, H.D., Moder, K., Thaler-Knoflach, B., Christian, E., Mendes, L.F., Schlick-Steiner, B.C. & Steiner, F.M. (2012) A toolbox for integrative species delimitation in Machilis jumping bristletails (Microcoryphia: Machilidae). *Zoologischer Anzeiger - A Journal of Comparative Zoology*, **251**, 307-316.
Deng, J., Wei, M., Yu, B. & Chen, Y. (2010) Efficient amplification of genes involved in microbial secondary metabolism by an improved genome walking method. *Applied Microbiology and Biotechnology*, **87**, 757-764.
Dobzhansky, T. (1940) Speciation as a Stage in Evolutionary Divergence. *The American Naturalist*, **74**, 312-321.
Dodt, M., Roehr, J., Ahmed, R. & Dieterich, C. (2012) FLEXBAR—Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology*, **1**, 895-905.
Domingues, V.S., Poh, Y.-P., Peterson, B.K., Pennings, P.S., Jensen, J.D. & Hoekstra, H.E. (2012) Evidence of adaptation from ancestral variation in young populations of beach mice. *Evolution*, **66**, 3209-3223.
Drummond, A. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology*, **7**, 214.
Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-1973.
Dun, G., Li, X., Cao, H., Zhou, R. & Li, L. (2007) Variations of Melanocortin Receptor 1 (MC1R) Gene in Three Pig Breeds. *Journal of Genetics and Genomics*, **34**, 777-782.
Dzierma, Y., Rabbel, W., Thorwart, M.M., Flueh, E.R., Mora, M.M. & Alvarado, G.E. (2011) The steeply subducting edge of the Cocos Ridge: Evidence from receiver functions beneath the northern Talamanca Range, south-central Costa Rica. *Geochemistry, Geophysics, Geosystems*, **12**, Q04S30, doi:10.1029/2010GC003477.
Echeverri, S., Cardona, A., Pardo, A., Monsalve, G., Valencia, V.A., Borrero, C., Rosero, S. & López, S. (2015) Regional provenance from southwestern Colombia fore-arc and intra-arc basins: implications for Middle to Late Miocene orogeny in the Northern Andes. *Terra Nova*, DOI: 10.1111/ter.12167.

Edwards, D.L. & Knowles, L.L. (2014) Species detection and individual assignment in species delimitation: can integrative data increase efficacy? *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **281**, 20132765.
Ellingson, R.A., Swift, C.C., Findley, L.T. & Jacobs, D.K. (2014) Convergent evolution of ecomorphological adaptations in geographically isolated Bay gobies (Teleostei: Gobionellidae) of the temperate North Pacific. *Molecular Phylogenetics and Evolution*, **70**, 464-477.
Endler, J.A. (1977) *Geographic variation, speciation, and clines*. Princeton University Press.
Endler, J.A. (1980) Natural Selection on Color Patterns in Poecilia reticulata. *Evolution*, **34**, 76-91.
Endler, J.A. (2012) A framework for analysing colour pattern geometry: adjacent colours. *Biological Journal of the Linnean Society*, **107**, 233-253.
Evans, M., Smith, S., Flynn, R. & Donoghue, M. (2009) Climate, niche evolution, and diversification of the "Bird - Cage" evening primroses (Oenothera, sections Anogra and Kleinia). *The American Naturalist*, **173**, 225-240.
Everts, R.E., Rothuizen, J. & van Oost, B.A. (2000) Identification of a premature stop codon in the melanocyte-stimulating hormone receptor gene (MC1R) in Labrador and Golden retrievers with yellow coat colour. *Animal Genetics*, **31**, 194-199.
Fitzpatrick, B.M. & Turelli, M. (2006) The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution*, **60**, 601-615.
Fontanesi, L., Beretti, F., Riggio, V., Dall'Olio, S., González, E.G., Finocchiaro, R., Davoli, R., Russo, V. & Portolano, B. (2009) Missense and nonsense mutations in melanocortin 1 receptor (MC1R) gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences. *BMC Genetics*, **10**, 1-12.
Forsman, A. & Shine, R. (1995) The adaptive significance of colour pattern polymorphism in the Australian scincid lizard Lampropholis delicata. *Biological Journal of the Linnean Society*, **55**, 273-291.
Fourcade, Y., Engler, J.O., Rödder, D. & Secondi, J. (2014) Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS ONE*, **9**, e97122.
Fraley, C. & Raftery, A.E. (2002) Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association*, **97**, 611-631.
Frost, S.K. & Robinson, S.J. (1984) Pigment cell differentiation in the fire-bellied toad, *Bombina orientalis*. I. Structural, chemical, and physical aspects of the adult pigment pattern. *Journal of Morphology*, **179**, 229-242.
Fu, Y.X. & Li, W.H. (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693-709.
Funk, V.A. & Richardson, K.S. (2002) Systematic Data in Biodiversity Studies: Use It or Lose It. *Systematic Biology*, **51**, 303-316.
Funk, W.C., Caminer, M. & Ron, S.R. (2011) High levels of cryptic species diversity uncovered in Amazonian frogs. *Proceedings of the Royal Society B: Biological Sciences*,
Funkhouser, J. (1956) New frogs from Ecuador and southwestern Colombia. *Zoologica*, **41**, 73-80.
Futuyma, D.J. & Mayer, G.C. (1980) Non-Allopatric Speciation in Animals. *Systematic Biology*, **29**, 254-271.

Galindo-Uribe, D., Sunyer, J., Hauswaldt, J.S., Amézquita, A., Pröhl, H. & Vences, M. (2014) Colour and pattern variation and Pleistocene phylogeographic origin of the strawberry poison frog, Oophaga pumilio, in Nicaragua. *Salamandra*, **50**, 225-235.

Gamberale-Stille, G. & Guilford, T. (2003) Contrast versus colour in aposematic signals. *Animal Behaviour*, **65**, 1021-1026.

Gangoso, L., Grande, J.M., Ducrest, A.L., Figuerola, J., Bortolotti, G.R., AndrÉS, J.A. & Roulin, A. (2011) MC1R-dependent, melanin-based colour polymorphism is associated with cell-mediated response in the Eleonora's falcon. *Journal of Evolutionary Biology*, **24**, 2055-2063.

Gould, B., McCouch, S. & Geber, M. (2015) *De Novo*Transcriptome Assembly and Identification of Gene Candidates for Rapid Evolution of Soil Al Tolerance in *Anthoxanthum odoratum* at the Long-Term Park Grass Experiment. *PLoS ONE*, **10**, e0124424.

Gould, S.J. & Johnston, R.F. (1972) Geographic Variation. *Annual Review of Ecology and Systematics*, **3**, 457-498.

Gräfe, K., Frisch, W., Villa, I.M. & Meschede, M. (2002) Geodynamic evolution of southern Costa Rica related to low-angle subduction of the Cocos Ridge: constraints from thermochronology. *Tectonophysics*, **348**, 187-204.

Graham, C.H., Ron, S.R., Santos, J.C., Schneider, C.J. & Moritz, C. (2004) Integrating phylogenetics and Environmental Niche Models to explore speciation mechanisms in Dendrobatid frogs. *Evolution*, **58**, 1781-1793.

Grant, T., Frost, D.R., Caldwell, J.P., Gagliardo, R.O.N., Haddad, C.F.B., Kok, P.J.R., Means, D.B., Noonan, B.P., Schargel, W.E. & Wheeler, W.C. (2006) Phylogenetic systematics of dart-poison frogs and their relatives (Amphibia: Athesphatanura: Dendrobatidae). *Bulletin of the American Museum of Natural History*, 1-262.

Gratten, J., Beraldi, D., Lowder, B.V., McRae, A.F., Visscher, P.M., Pemberton, J.M. & Slate, J. (2007) Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 619-626.

Greenwood, J.J.D., Wood, E.M. & Batchelor, S. (1981) Apostatic selection of distasteful prey. *Heredity*, **47**, 27-34.

Griffith, S.C., Parker, T.H. & Olson, V.A. (2006) Melanin- versus carotenoid-based sexual signals: is the difference really so black and red? *Animal Behaviour*, **71**, 749-763.

Gross, J.B., Borowsky, R. & Tabin, C.J. (2009) A novel role for MC1R in the parallel evolution of depigmentation in independent populations of the Cavefish *Astyanax mexicanus*. *PLoS Genet*, **5**, e1000326.

Grummer, J.A., Bryson, R.W. & Reeder, T.W. (2013) Species Delimitation Using Bayes Factors: Simulations and Application to the Sceloporus scalaris Species Group (Squamata: Phrynosomatidae). *Systematic Biology*,

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N. & Regev, A. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protocols*, **8**, 1494-1512.

Haffer, J. (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131-137.

Haffer, J.r. (1997) Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiversity and Conservation*, **6**, 451-476.
Hagemann, S. & Pröhl, H. (2007) Mitochondrial paraphyly in a polymorphic poison frog species (Dendrobatidae; *D. pumilio*). *Molecular Phylogenetics and Evolution*, **45**, 740-747.
Hagman, M. & Forsman, A. (2003) Correlated evolution of conspicuous coloration and body size in poison frogs (Dendrobatidae). *Evolution*, **57**, 2904-2910.
Hammer, Ø., Harper, D. & Ryan, P. (2001) PAST: Paleontological Statistics Software Package for education and data analysis. Palaeontolia Electronica. **4(1)**, 9.
Harmon, L.J., Schulte, J.A., Larson, A. & Losos, J.B. (2003) Tempo and mode of evolutionary radiation in Iguanian lizards. *Science*, **301**, 961-964.
Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. (2008) GEIGER: investigating evolutionary radiations. *Bioinformatics*, **24**, 129-131.
Hartmann, T. (1999) Chemical ecology of pyrrolizidine alkaloids. *Planta*, **207**, 483-495.
Harvey, P.H., Bull, J.J., Pemberton, M. & Paxton, R.J. (1982) The Evolution of Aposematic Coloration in Distasteful Prey: A Family Model. *The American Naturalist*, **119**, 710-719.
Hashimoto, T. & Yamada, Y. (1994) Alkaloid biogenesis: molecular aspects. *Annual review of plant biology*, **45**, 257-285.
Hausdorf, B. & Hennig, C. (2010) Species Delimitation Using Dominant and Codominant Multilocus Markers. *Systematic Biology*,
Hauswaldt, J.S., Ludewig, A.-K., Vences, M. & Pröhl, H. (2011) Widespread co-occurrence of divergent mitochondrial haplotype lineages in a Central American species of poison frog (Oophaga pumilio). *Journal of Biogeography*, **38**, 711-726.
Hauswaldt, J.S., Ludewig, A.-K., Hagemann, S., Pröhl, H. & Vences, M. (2009) Ten microsatellite loci for the strawberry poison frog (Oophaga pumilio). *Conservation Genetics*, **10**, 1935-1937.
Hedrick, P.W. (1995) Gene Flow and Genetic Restoration: The Florida Panther as a Case Study. *Conservation Biology*, **9**, 996-1007.
Hegna, R., Saporito, R. & Donnelly, M. (2012) Not all colors are equal: predation and color polytypism in the aposematic poison frog *Oophaga pumilio*. *Evolutionary Ecology*, 1-15.
Hegna, R.H., Saporito, R.A., Gerow, K.G. & Donnelly, M.A. (2011) Contrasting colors of an aposematic poison frog do not affect predation. *Annales Zoologici Fennici*, **48**, 29-38.
Heibl, C. & Calenge, C. (2013) Integrating phylogenetics and climatic niche modeling. In: *http://cran.r-project.org/web/packages/phyloclim/phyloclim.pdf*
Helbig, A.J., Knox, A.G., Parkin, D.T., Sangster, G. & Collinson, M. (2002) Guidelines for assigning species rank†. *Ibis*, **144**, 518-525.
Heled, J. & Drummond, A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570-580.
Herczeg, G., Matsuba, C. & Merilä, J. (2010) Sequence variation in the melanocortin-1 receptor gene (Mc1r) does not explain variation in the degree of melanism in a widespread amphibian. *Ann. Zool. Fennici*, **47**, 37-45.
Hijmans, R., Guarino, L., Cruz, M. & Rojas, E. (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*, 15-19.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978.

Hoekstra, H.E. (2006) Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity*, **97**, 222-234.
Hoekstra, H.E., Hirschmann, R.J., Bundey, R.A., Insel, P.A. & Crossland, J.P. (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science*, **313**, 101-104.
Hoffman, E.A. & Blouin, M.S. (2000) A review of colour and pattern polymorphisms in anurans. *Biological Journal of the Linnean Society*, **70**, 633-665.
Holen, Ø.H. (2013) Disentangling taste and toxicity in aposematic prey. *Proceedings of the Royal Society B: Biological Sciences*, **280**
Holen, Ø.H. & Svennungsen, T.O. (2012) Aposematism and the Handicap Principle. *The American Naturalist*, **180**, 629-641.
Howlett, R. (2000) Hotspots: Earth's biologically richest and most endangered terrestrial ecoregions. *Nature*, **406**, 237-238.
Hubert, N. & Renno, J.-F. (2006) Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, **33**, 1414-1436.
Hubisz, M.J., Falush, D., Stephens, M. & Pritchard, J.K. (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322-1332.
Huelsenbeck, J.P., Andolfatto, P. & Huelsenbeck, E.T. (2011) Structurama: Bayesian Inference of Population Structure. *Evolutionary Bioinformatics Online*, **7**, 55-59.
Huxley, J. (1958) *A Book that shook the world; anniversary essays on Charles Darwin's Origin of species*. University of Pittsburgh Press, Pittsburgh.
Ihaka, R. & Gentleman, R. (1996) R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, **5**, 299-314.
Illoldi-Rangel, P., Sanchez-Cordero, V. & Townsend Peterson, A. (2004) Predicting distributions of Mexican mammals using Ecological Niche Modeling. *Journal of Mammalogy*, **85**, 658-662.
IUCN (2012) *IUCN Red List of Threatened Species. Version 2012.1,*. Available at: http://www.iucnredlist.org. (accessed December 2013).
Jang, Y., Hahm, E.H., Lee, H.-J., Park, S., Won, Y.-J. & Choe, J.C. (2011) Geographic Variation in Advertisement Calls in a Tree Frog Species: Gene Flow and Selection Hypotheses. *PLoS ONE*, **6**, e23297.
Joly, S., Heenan, P.B. & Lockhart, P.J. (2014) Species radiation by niche shifts in New Zealand's rockcresses (Pachycladon, Brassicaceae). *Systematic Biology*, **63**, 192-202.
Jombart, T., Devillard, S. & Balloux, F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.
Kaeuffer, R., Peichel, C.L., Bolnick, D.I. & Hendry, A.P. (2012) Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution*, **66**, 402-418.
Katoh, K. & Standley, D.M. (2014) MAFFT: Iterative Refinement and Additional Methods. *Multiple Sequence Alignment Methods* (ed. by J.D. Russell), pp. 131-146. Humana Press, Totowa, NJ.
Katoh, K., Misawa, K., Kuma, K.i. & Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059-3066.

Klovins, J., Haitina, T., Fridmanis, D., Kilianova, Z., Kapa, I., Fredriksson, R., Gallo-Payet, N. & Schiöth, H.B. (2004) The Melanocortin System in Fugu: Determination of POMC/AGRP/MCR Gene Repertoire and Synteny, As Well As Pharmacology and Anatomical Distribution of the MCRs. *Molecular Biology and Evolution*, **21**, 563-579.
Klungland, H., Våge, D.I., Gomez-Raya, L., Adalsteinsson, S. & Lien, S. (1995) The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mamm Genome*, **6**
Klungland, H., Røed, K.H., Nesbø, C.L., Jakobsen, K.S. & Våge, D.I. (1999) The melanocyte-stimulating hormone receptor (MC1-R) gene as a tool in evolutionary studies of Artiodactyles. *Hereditas*, **131**
Knowlton, N. & Weigt, L.A. (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 2257-2263.
Kornobis, E., Cabellos, L., Aguilar, F., Frías-López, C., Rozas, J., Marco, J. & Zardoya, R. (2015) TRUFA: A User-Friendly Web Server for de novo RNA-seq Analysis Using Cluster Computing. *Evolutionary Bioinformatics Online*, **11**, 97-104.
Kraemer, A.C., Kissner, J. & Adams, D.C. (2012) Morphological color-change in the Red-Backed salamander (*Plethodon cinereus*) while kept in captivity. *Copeia*, **2012**, 748-755.
Langham, G.M. & Benkman, C. (2004) Specialized avian predators repeatedly attack novel color morphs of Heliconius butterflies. *Evolution*, **58**, 2783-2787.
Lawson, A.M. & Weir, J.T. (2014) Latitudinal gradients in climatic-niche evolution accelerate trait evolution at high latitudes. *Ecology Letters*, **17**, 1427-1436.
Leaché, A.D. & Fujita, M.K. (2010) Bayesian species delimitation in West African forest geckos (Hemidactylus fasciatus). *Proceedings of the Royal Society of London, Series B: Biological Sciences*,
Leaché, A.D., Fujita, M.K., Minin, V.N. & Bouckaert, R.R. (2014) Species Delimitation using Genome-Wide SNP Data. *Systematic Biology*, **63**, 534-542.
Li, L., Stoeckert, C.J. & Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome research*, **13**, 2178-2189.
Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451-1452.
Logan, D.W., Burn, S.F. & Jackson, I.J. (2006) Regulation of pigmentation in zebrafish melanophores. *Pigment Cell Research*, **19**, 206-213.
Lötters, S. (1992) Zur Validitat von *Dendrobates lehmani* Myers & Daly, 1976 aufgrund zweier neuer Farbformen von *Dendrobates histrionicus* Berthold, 1845. *Salamandra*, **15**, 138-144.
Lötters, S., Glaw, F., Kohler, F. & Castro, F. (1999) On the geographic variation of the advertisement call of *Dendrobates histrionicus* Berthold, 1845 and related forms from north-western South America. *Herpetozoa*, **12**, 23-38.
Lougheed, S., Austin, J., Bogart, J., Boag, P. & Chek, A. (2006) Multi-character perspectives on the evolution of intraspecific differentiation in a neotropical hylid frog. *Bmc Evolutionary Biology*, **6**, 23.
Lynch, J.D. & Duellman, W.E. (1997) *Frogs of the genus Eleutherodactylus (Leptodactylidae) in western Ecuador: systematics, ecology, and biogeography*. Lawrence, Kan.: Natural History Museum, University of Kansas.

Lynch, J.D. & Arroyo, S.B. (2009) Risks to Colombian Amphibian Fauna from Cultivation of Coca (*Erythroxylum coca*): A Geographical Analysis. *Journal of Toxicology and Environmental Health, Part A*, **72**, 974-985.
Lynn, S.K. (2005) Learning to avoid aposematic prey. *Animal Behaviour*, **70**, 1221-1226.
Maan, M.E. & Cummings, M.E. (2008) Female preferences for aposematic signal components in a polymorphic poison frog. *Evolution*, **62**, 2334-2345.
Maan, M.E. & Cummings, M.E. (2009) Sexual dimorphism and directional sexual selection on aposematic signals in a poison frog. *Proceedings of the National Academy of Sciences*, **106**, 19072-19077.
Macey, J.R., Schulte Ii, J.A., Larson, A., Fang, Z., Wang, Y., Tuniyev, B.S. & Papenfuss, T.J. (1998) Phylogenetic relationships of toads in the *Bufo bufo* species group from the eastern escarpment of the Tibetan plateau: a case of vicariance and dispersal. *Molecular Phylogenetics and Evolution*, **9**, 80-87.
Maddison, W.P. & Knowles, L.L. (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, **55**, 21-30.
Mallet, J. & Joron, M. (1999) Evolution of diversity in warning color and mimicry: Polymorphisms, Shifting Balance, and Speciation. *Annual Review of Ecology and Systematics*, **30**, 201-233.
Mangano, F.T., Fukuzawa, T., Johnson, W.C. & Bagnara, J.T. (1992) Intrinsic pigment cell stimulating activity in the skin of the leopard frog, *Rana pipiens*. *Journal of Experimental Zoology*, **263**, 112-118.
Mappes, J., Marples, N. & Endler, J.A. (2005) The complex business of survival by aposematism. *Trends in Ecology & Evolution*, **20**, 598-603.
Mayaux, P., Achard, F. & Malingreau, J.P. (1998) Global tropical forest area measurements derived from coarse resolution satellite imagery: a comparison with other approaches. *Environmental Conservation*, **25**, 37-52.
Mayr, E. (1942) *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press.
Mayr, E. (1963) *Animal species and evolution*. Harvard University Press; London: Oxford University Press.
McCormack, J.E., Zellmer, A.J. & Knowles, L.L. (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: insights from tests with niche models. *Evolution*, **64**, 1231-1244.
McDonald, J.H. & Kreitman, M. (1991) Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, **351**, 652-654.
Medina, I., Wang, I.J., Salazar, C. & Amézquita, A. (2013) Hybridization promotes color polymorphism in the aposematic harlequin poison frog, *Oophaga histrionica*. *Ecology and Evolution*, **3**, 4388-4400.
Mendez-Narvaez, J. & Amezquita, A. (2014) Physical combat in the poison-arrow frog, Kokoé-pá (*Oophaga histrionica*) from Arusi, Choco, Colombia. *Herpetology Notes*, **7**, 1-2.
Merow, C., Smith, M.J. & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, **36**, 1058-1069.
Micheletti, S., Parra, E. & Routman, E.J. (2012) Adaptive color polymorphism and unusually high local genetic diversity in the Side-Blotched lizard, *Uta stansburiana*. *PLoS ONE*, **7**, e47694.

Mills, M.G. & Patterson, L.B. (2009) Not just black and white: Pigment pattern development and evolution in vertebrates. *Seminars in Cell & Developmental Biology*, **20**, 72-81.
Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D.F. & Wright, F. (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, **25**, 126-127.
Montes, C., Cardona, A., McFadden, R., Morón, S.E., Silva, C.A., Restrepo-Moreno, S., Ramírez, D.A., Hoyos, N., Wilson, J., Farris, D., Bayona, G.A., Jaramillo, C.A., Valencia, V., Bryan, J. & Flores, J.A. (2012) Evidence for middle Eocene and younger land emergence in central Panama: Implications for Isthmus closure. *Geological Society of America Bulletin*, **124**, 780-799.
Moritz, C., Patton, J.L., Schneider, C.J. & Smith, T.B. (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533-563.
Mundy, N.I. (2009) Conservation and Convergence of Colour Genetics: MC1R Mutations in brown Cavefish. *PLoS Genet*, **5**, e1000388.
Myers, C.E. (1982) Spotted poison frogs: Description of Three new Dendrobates from Western Amazonia, and resurrection of a lost species from "Chiriqui". *American Museum Novitates*, **2721**, 1-23.
Myers, C.W. & Daly, J.W. (1976) Preliminary evaluation of skin toxins and vocalizations in taxonomic and evolutionary studies of poison-dart frogs (Dendrobatidae). *Bulletin of the American Museum of Natural History*, 175-262.
Nakazato, T., Warren, D.L. & Moyle, L.C. (2010) Ecological and geographic modes of species divergence in wild tomatoes. *American Journal of Botany*, **97**, 680-693.
Newton, J.M., Wilkie, A.L., He, L., Jordan, S.A., Metallinos, D.L., Holmes, N.G., Jackson, I.J. & Barsh, G.S. (2000) Melanocortin 1 receptor variation in the domestic dog. *Mamm Genome*, **11**
NOAA (2013) *Global Measured Extremes of Temperature and Precipitation*. Available at: http://www.ncdc.noaa.gov/oa/climate/globalextremes.html#highpre (accessed April 16 2013 2013).
Nokelainen, O., Hegna, R.H., Reudler, J.H., Lindstedt, C. & Mappes, J. (2012) Trade-off between warning signal efficacy and mating success in the wood tiger moth. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 257-265.
Noonan, B.P. & Comeault, A.A. (2009) The role of predator selection on polymorphic aposematic poison frogs. *Biology Letters*, **5**, 51-54.
Nowakowski, A.J., Otero Jiménez, B., Allen, M., Diaz-Escobar, M. & Donnelly, M.A. (2013) Landscape resistance to movement of the poison frog, *Oophaga pumilio*, in the lowlands of northeastern Costa Rica. *Animal Conservation*, **16**, 188-197.
Ohbayashi, N., Maruta, Y., Ishida, M. & Fukuda, M. (2012) Melanoregulin regulates retrograde melanosome transport through interaction with the RILP–p150Glued complex in melanocytes. *Journal of Cell Science*, **125**, 1508-1518.
Oksanen, J., Blanchet, F., Kindt, R., Legendre, P. & Minchin, P. (2013) vegan: Community Ecology Package. R package version 2.0-6. *Available at: http://mirror.bjtu.edu.cn/cran/web/packages/vegan/*,
Otani, M., Shitan, N., Sakai, K., Martinoia, E., Sato, F. & Yazaki, K. (2005) Characterization of Vacuolar Transport of the Endogenous Alkaloid Berberine in Coptis japonica. *Plant Physiology*, **138**, 1939-1946.

Padial, J.M. & De la Riva, I. (2006) Taxonomic Inflation and the Stability of Species Lists: The Perils of Ostrich's Behavior. *Systematic Biology*, **55**, 859-867.
Padial, J.M. & De La Riva, I. (2009) Integrative taxonomy reveals cryptic Amazonian species of Pristimantis (Anura: Strabomantidae). *Zoological Journal of the Linnean Society*, **155**, 97-122.
Padial, J.M. & De La Riva, I. (2010) A response to recent proposals for integrative taxonomy. *Biological Journal of the Linnean Society*, **101**, 747-756.
Padial, J.M., Miralles, A., De la Riva, I. & Vences, M. (2010) The integrative future of taxonomy. *Frontiers in Zoology*, **7**, 1-14.
Palumbo, A., D Cosmo, A., Gesualdo, I. & Vincent, J. (1997) Subcellular localization and function of melanogenic enzymes in the ink gland of Sepia officinalis. *Biochemical Journal*, **323**, 749-756.
Parra, G., Bradnam, K. & Korf, I. (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*, **23**, 1061-1067.
Parra, G., Bradnam, K., Ning, Z., Keane, T. & Korf, I. (2009) Assessing the gene space in draft genomes. *Nucleic Acids Research*, **37**, 289-297.
Paulay, G. & Meyer, C. (2002) Diversification in the tropical pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Integrative and Comparative Biology*, **42**, 922-934.
Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, **28**, 2537-2539.
Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H. & Stockwell, D.R.B. (2002) Future projections for Mexican faunas under global climate change scenarios. *Nature*, **416**, 626-629.
Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.
Phillott, A.D., McDonald, K.R. & Skerratt, L.F. (2011) Inflammation in digits of unmarked and toe-tipped wild hylids. *Wildlife Research*, **38**, 204-207.
Posada, D. (2008) jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, **25**, 1253-1256.
Posada, D. & Buckley, T.R. (2004) Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. *Systematic Biology*, **53**, 793-808.
Posso-Terranova, A. & Andrés, J. (2016a) Complex niche divergence underlies lineage diversification in *Oophaga* poison frogs. *Journal of Biogeography*, **DOI: 10.1111/jbi.12799**
Posso-Terranova, A. & Andrés, J. (2016b) Ecology, molecules and colour: Multivariate species delimitation and conservation of Harlequin poison frogs. *bioRxiv*, doi: http://dx.doi.org/10.1101/050922.
Poulton, E.B. (1887) The Experimental Proof of the Protective Value of Colour and Markings in Insects in reference to their Vertebrate Enemies. *Proceedings of the Zoological Society of London*, **55**, 191-274.
Poulton, E.B. (1890) *The colours of animals*, 2d. edn. Kegan Paul, Trench Tribner, London.
Price, T.D., Hooper, D.M., Buchanan, C.D., Johansson, U.S., Tietze, D.T., Alstrom, P., Olsson, U., Ghosh-Harihar, M., Ishtiaq, F., Gupta, S.K., Martens, J., Harr, B., Singh, P. &

Mohan, D. (2014) Niche filling slows the diversification of Himalayan songbirds. *Nature*, **509**, 222-225.
Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945-959.
Pröhl, H. & Ostrowski, T. (2011) Behavioural elements reflect phenotypic colour divergence in a poison frog. *Evolutionary Ecology*, **25**, 993-1015.
Protas, M.E. & Patel, N.H. (2008) Evolution of Coloration Patterns. *Annual Review of Cell and Developmental Biology*, **24**, 425-446.
Prudic, K.L., Skemp, A.K. & Papaj, D.R. (2007) Aposematic coloration, luminance contrast, and the benefits of conspicuousness. *Behavioral Ecology*, **18**, 41-46.
Puchtler, H. & Meloan, S.N. (1985) On the chemistry of formaldehyde fixation and its effects on immunohistochemical reactions. *Histochemistry*, **82**, 201-204.
Pyron, R.A. & Wiens, J.J. (2013) Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proceedings of the Royal Society B: Biological Sciences*, **280**, DOI: 10.1098/rspb.2013.1622.
Qvarnström, A., Rudh, A., Edström, T., Ödeen, A., Løvlie, H. & Tullberg, B.S. (2014) Coarse dark patterning functionally constrains adaptive shifts from aposematism to crypsis in strawberry poison frogs. *Evolution*, **68**, 2793-2803.
Radosavljevic, A. & Anderson, R.P. (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography*, **41**, 629-643.
Ramos, R.R. & Freitas, A. (1999) Population biology and wing color variation in Heliconius erato phyllis (Nymphalidae). *Journal of the Lepidopterists Society*, **53**, 11-21.
Rand, D.M. & Kann, L.M. (1996) Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from Drosophila, mice, and humans. *Molecular Biology and Evolution*, **13**, 735-748.
Ratcliffe, J.M. & Nydam, M.L. (2008) Multimodal warning signals for a multiple predator world. *Nature*, **455**, 96-99.
Rees, J.L. (2000) The Melanocortin 1 Receptor (MC1R): More Than Just Red Hair. *Pigment Cell Research*, **13**, 135-140.
Reeves, P.A. & Richards, C.M. (2011) Species delimitation under the general lineage concept: an empirical example using wild North American hops (Cannabaceae: *Humulus lupulus*). *Systematic Biology*, **60**, 45-59.
Richards-Zawacki, C.L. & Cummings, M.E. (2011) Intraspecific reproductive character displacement in a polymorphic poison frog *Dendrobates pumilio*. *Evolution*, **65**, 259-267.
Richards-Zawacki, C.L., Wang, I.J. & Summers, K. (2012) Mate choice and the genetic basis for colour variation in a polymorphic dart frog: inferences from a wild pedigree. *Molecular Ecology*, **21**, 3879-3892.
Robbins, L.S., Nadeau, J.H., Johnson, K.R., Kelly, M.A., Roselli-Rehfuss, L., Baack, E., Mountjoy, K.G. & Cone, R.D. (1993) Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell*, **72**
Roberts, A. & Pachter, L. (2013) Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat Meth*, **10**, 71-73.
Roberts, J.L., Brown, J.L., May, R., Arizabal, W., Schulte, R. & Summers, K. (2006) Genetic divergence and speciation in lowland and montane peruvian poison frogs. *Molecular Phylogenetics and Evolution*, **41**, 149-64.

Rojas, B. (2012) *The apparent paradox of colour variation in aposematic poison frogs*. Dissertation, Deakin University, Australia.
Rojas, B. (2016) Behavioural, ecological, and evolutionary aspects of diversity in frog colour patterns. *Biological Reviews*, n/a-n/a.
Rosenblum, E.B., Hoekstra, H.E. & Nachman, M.W. (2004) Adaptive reptile color variation and the evolution of the MCIR gene. *Evolution*, **58**, 1794-1808.
Rosenblum, E.B., Römpler, H., Schöneberg, T. & Hoekstra, H.E. (2010) Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proceedings of the National Academy of Sciences*, **107**, 2113-2117.
Ruane, S., Bryson, R.W., Pyron, R.A. & Burbrink, F.T. (2014) Coalescent Species Delimitation in Milksnakes (Genus *Lampropeltis*) and Impacts on Phylogenetic Comparative Analyses. *Systematic Biology*, **63**, 231-250.
Rudh, A. & Qvarnström, A. (2013) Adaptive colouration in amphibians. *Seminars in Cell & Developmental Biology*,
Rudh, A., Rogell, B. & Hoglund, J. (2007) Non-gradual variation in colour morphs of the strawberry poison frog Dendrobates pumilio: genetic and geographical isolation suggest a role for selection in maintaining polymorphism. *Molecular Ecology*, **16**, 4284-4294.
Ruxton, G.D., Sherratt, T.N. & Speed, M.P. (2004) The form and function of warning displays. In. Oxford University Press, Oxford.
Sakai, K., Shitan, N., Sato, F., Ueda, K. & Yazaki, K. (2002) Characterization of berberine transport into Coptis japonica cells and the involvement of ABC protein. *Journal of Experimental Botany*, **53**, 1879-1886.
Santos, J.C., Coloma, L.A. & Cannatella, D.C. (2003) Multiple, recurring origins of aposematism and diet specialization in poison frogs. *Proceedings of the National Academy of Sciences*, **100**, 12792-12797.
Santos, J.C., Tarvin, R.D. & O'Connell, L.A. (2016) A Review of Chemical Defense in Poison Frogs (Dendrobatidae): Ecology, Pharmacokinetics, and Autoresistance. *Chemical Signals in Vertebrates 13* (ed. by A.B. Schulte, E.T. Goodwin and H.M. Ferkin), pp. 305-337. Springer International Publishing, Cham.
Santos, J.C., Coloma, L.A., Summers, K., Caldwell, J.P., Ree, R. & Cannatella, D.C. (2009) Amazonian Amphibian Diversity Is Primarily Derived from Late Miocene Andean Lineages. *PLoS Biol*, **7**, e1000056.
Saporito, R., Donnelly, M., Spande, T. & Garraffo, H.M. (2012) A review of chemical ecology in poison frogs. *Chemoecology*, **22**, 159-168.
Saporito, R., Donnelly, M., Garraffo, H., Spande, T. & Daly, J. (2006) Geographic and seasonal variation in alkaloid-based chemical defenses of *Dendrobates pumilio* from Bocas del Toro, Panama. *Journal of Chemical Ecology*, **32**, 795-814.
Saporito, R.A., Garraffo, H.M., Donnelly, M.A., Edwards, A.L., Longino, J.T. & Daly, J.W. (2004) Formicine ants: An arthropod source for the pumiliotoxin alkaloids of dendrobatid poison frogs. *Proceedings of the National Academy of Sciences*, **101**, 8045-8050.
Saporito, R.A., Donnelly, M.A., Jain, P., Martin Garraffo, H., Spande, T.F. & Daly, J.W. (2007) Spatial and temporal patterns of alkaloid variation in the poison frog *Oophaga pumilio* in Costa Rica and Panama over 30 years. *Toxicon*, **50**, 757-778.
Schluter, D. (2009) Evidence for ecological speciation and its alternative. *Science*, **323**, 737-741.
Schluter, D., Price, T., Mooers, A.Ø. & Ludwig, D. (1997) Likelihood of ancestor states in adaptive radiation. *Evolution*, **51(6)**, 1699-1711.

Schulte Ii, J.A., Macey, J.R., Larson, A. & Papenfuss, T.J. (1998) Molecular tests of phylogenetic taxonomies: a general procedure and example using four subfamilies of the lizard family Iguanidae. *Molecular Phylogenetics and Evolution*, **10**, 367-376.
Searle, A.G. (1968) *Comparative Genetics of Coat Colour in Mammals*. Logos Press, London.
Selz, Y., Braasch, I., Hoffmann, C., Schmidt, C., Schultheis, C., Schartl, M. & Volff, J.-N. (2007) Evolution of melanocortin receptors in teleost fish: The melanocortin type 1 receptor. *Gene*, **401**, 114-122.
Sheffield, J. (2007) ImageJ: A useful tool for biological image processing and analysis. *Microscopy and Microanalysis*, **13**, 200-201.
Sherratt, T. (2008) The evolution of Müllerian mimicry. *Naturwissenschaften*, **95**, 681-695.
Sherratt, T.N. (2006) Spatial mosaic formation through frequency-dependent selection in Müllerian mimicry complexes. *Journal of Theoretical Biology*, **240**, 165-174.
Shitan, N. & Yazaki, K. (2007) Accumulation and Membrane Transport of Plant Alkaloids. *Current Pharmaceutical Biotechnology*, **8**, 244-252.
Shitan, N., Bazin, I., Dan, K., Obata, K., Kigawa, K., Ueda, K., Sato, F., Forestier, C. & Yazaki, K. (2003) Involvement of CjMDR1, a plant multidrug-resistance-type ATP-binding cassette protein, in alkaloid transport in Coptis japonica. *Proceedings of the National Academy of Sciences*, **100**, 751-756.
Sigel, A., Sigel, H. & Sigel, R.K. (2007) *The ubiquitous roles of cytochrome P450 proteins: metal ions in life sciences*. John Wiley & Sons.
Silverman, B.W. (1986) *Density estimation for statistics and data analysis*. CRC press.
Silvers, W.K. (1979) *The coat colors of mice: a model for mammalian gene action and interaction*. Springer -Verlag, New York.
Silverstone, P.A. (1973) Observations on the Behavior and Ecology of a Colombian Poison-Arrow Frog, the Kõkoé-Pá (Dendrobates histrionicus Berthold). *Herpetologica*, **29**, 295-301.
Silverstone, P.A. (1975) A revision of the poison-arrow frogs of the genus *Dendrobates* Wagler. *Natural History Museum of Los Angeles County*, **21**, 1-55.
Smith, B.T., McCormack, J.E., Cuervo, A.M., Hickerson, M.J., Aleixo, A., Cadena, C.D., Perez-Eman, J., Burney, C.W., Xie, X., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Derryberry, E.P., Prejean, J., Fields, S. & Brumfield, R.T. (2014) The drivers of tropical speciation. *Nature*, **515**, 406-409.
Smith, J.M. (1966) Sympatric Speciation. *The American Naturalist*, **100**, 637-650.
Snezhko, A., Barlan, K., Aranson, I.S. & Gelfand, V.I. (2010) Statistics of Active Transport in Xenopus Melanophores Cells. *Biophysical Journal*, **99**, 3216-3223.
Soberón, J. & Peterson, A.T. (2011) Ecological niche shifts and environmental space anisotropy: a cautionary note. *Revista Mexicana de Biodiversidad*, **82**, 1384-1355.
Solomon, S.E., Bacci, M., Jr., Martins, J., Jr., Vinha, G.G. & Mueller, U.G. (2008) Paleodistributions and comparative molecular phylogeography of Leafcutter Ants (*Atta* spp.) provide new insight into the origins of Amazonian diversity. *PLoS ONE*, **3**, e2738.
Steiner, C.C., Römpler, H., Boettger, L.M., Schöneberg, T. & Hoekstra, H.E. (2009) The Genetic Basis of Phenotypic Convergence in Beach Mice: Similar Pigment Patterns but Different Genes. *Molecular Biology and Evolution*, **26**, 35-45.
Stephens, M., Smith, N.J. & Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*, **68**

Stevens, M., Parraga, C.A., Cuthill, I.C., Partridge, J.C. & Troscianko, T.S. (2007) Using digital photography to study animal coloration. *Biological Journal of the Linnean Society*, **90**, 211-237.

Stothard, P. (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques*, **28**, 1102, 1104-1102, 1104.

Struwe, L., Smouse, P.E., Heiberg, E., Haag, S. & Lathrop, R.G. (2011) Spatial evolutionary and ecological vicariance analysis (SEEVA), a novel approach to biogeography and speciation research, with an example from Brazilian Gentianaceae. *Journal of Biogeography*, **38**, 1841-1854.

Sugimoto, M. (2002) Morphological color changes in fish: Regulation of pigment cell density and morphology. *Microscopy Research and Technique*, **58**, 496-503.

Summers, K. (2003) Convergent evolution of bright coloration and toxicity in frogs. *Proceedings of the National Academy of Sciences*, **100**, 12533-12534.

Summers, K., Cronin, T.W. & Kennedy, T. (2003) Variation in spectral reflectance among populations of *Dendrobates pumilio,* the strawberry poison frog, in the Bocas del Toro Archipelago, Panama. *Journal of Biogeography*, **30**, 35-53.

Summers, K., Bermingham, E., Weigt, L., McCafferty, S. & Dahistrom, L. (1997) Phenotypic and genetic divergence in three species of dart-poison frogs with contrasting parental behavior. *Journal of Heredity*, **88**, 8-13.

Svádová, K., Exnerová, A., Štys, P., Landová, E., Valenta, J., Fučíková, A. & Socha, R. (2009) Role of different colours of aposematic insects in learning, memory and generalization of naïve bird predators. *Animal Behaviour*, **77**, 327-336.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**, 2725-2729.

Tarvin, R.D., Santos, J.C., O'Connell, L.A., Zakon, H.H. & Cannatella, D.C. (2016) Convergent Substitutions in a Sodium Channel Suggest Multiple Origins of Toxin Resistance in Poison Frogs. *Molecular Biology and Evolution*,

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.

Våge, D.I., Klungland, H., Lu, D. & Cone, R.D. (1999) Molecular and pharmacological characterization of dominant black coat color in sheep. *Mamm Genome*, **10**

Valcárcel, V. & Vargas, P. (2010) Quantitative morphology and species delimitation under the general lineage concept: Optimization for Hedera (Araliaceae). *American Journal of Botany*, **97**, 1555-1573.

Valencia-Zuleta, A., Jaramillo-Martínez, A., Echeverry-Bocanegra, A., Viáfara-Vega, R., Hernández-Córdoba, O., Cardona-Botero, V., Gutiérrez-Zúñiga, J. & Castro-Herrera, F. (2014) Conservation status of the herpetofauna, protected areas, and current problems in Valle del Cauca, Colombia. *Amphibian & Reptile Conservation*, **8**, 1-18.

Valkonen, J.K., Nokelainen, O., Niskanen, M., Kilpimaa, J., Björklund, M. & Mappes, J. (2012) Variation in predator species abundance can cause variable selection pressure on warning signaling prey. *Ecology and Evolution*, **2**, 1971-1976.

Van Bel, M., Proost, S., Wischnitzki, E., Movahedi, S., Scheerlinck, C., Van de Peer, Y. & Vandepoele, K. (2012) Dissecting Plant Genomes with the PLAZA Comparative Genomics Platform. *Plant Physiology*, **158**, 590-600.

Van Gele, M., Dynoodt, P. & Lambert, J. (2009) Griscelli syndrome: a model system to study vesicular trafficking. *Pigment Cell & Melanoma Research*, **22**, 268-282.

Van Valen, L. (1976) Ecological species, multispecies, and oaks. *Taxon*, 233-239.

Vargas-Salinas, F. & Amézquita, A. (2013) Stream Noise, Hybridization, and Uncoupled Evolution of Call Traits in Two Lineages of Poison Frogs: *Oophaga histrionica* and *Oophaga lehmanni*. *PLoS ONE*, **8**, e77545.

Vences, M., Kosuch, J., Boistel, R., Haddad, C.F.B., La Marca, E., Lotters, S. & Veith, M. (2003) Convergent evolution of aposematic coloration in Neotropical poison frogs: a molecular phylogenetic perspective. *Organisms Diversity & Evolution*, **3**, 215-226.

Vieites, D.R., Wollenberg, K.C., Andreone, F., Köhler, J., Glaw, F. & Vences, M. (2009) Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proceedings of the National Academy of Sciences of the United States of America Biological Sciences*, **106**, 8267-8272.

Wagner, G.P., Kin, K. & Lynch, V.J. (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences*, **131**, 281-285.

Walther, C.H. (2003) The crustal structure of the Cocos ridge off Costa Rica. *Journal of Geophysical Research: Solid Earth (1978–2012)*, **108**

Wang, I.J. (2011) Inversely related aposematic traits: reduced conspicuousness evolves with increased toxicity in a polymorphic poison-dart frog. *Evolution*, **65**, 1637-1649.

Wang, I.J. & Shaffer, H.B. (2008) Rapid color evolution in an aposematic species: a phylogenetic analysis of color variation in the strikingly polymorphic strawberry poison-dart frog. *Evolution*, **62**, 2742-2759.

Wang, I.J., Glor, R.E. & Losos, J.B. (2013) Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecology Letters*, **16**, 175-182.

Warren, D.L. (2012) In defense of 'niche modeling'. *Trends in Ecology & Evolution*, **27**, 497-500.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883.

Warren, D.L., Glor, R.E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607-611.

Weir, J., T. & Price, T.D. (2011) Limits to speciation inferred from times to secondary sympatry and ages of hybridizing species along a latitudinal gradient. *The American Naturalist*, **177**, 462-469.

Willink, B., Brenes-Mora, E., Bolaños, F. & Pröhl, H. (2013) Not everything is black and white: color and behavioral variation reveal a continuum between cryptic and aposematic strategies in a polymorphic poison frog. *Evolution*, DOI: 10.1111/evo.12153.

Wise, H. (2012) The roles played by highly truncated splice variants of G protein-coupled receptors. *Journal of Molecular Signaling*, **7**, 13.

Wu, C.-I. (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851-865.

Wüster, W., Allum, C.S.E., Bjargardóttir, I.B., Bailey, K.L., Dawson, K.J., Guenioui, J., Lewis, J., McGurk, J., Moore, A.G., Niskanen, M. & Pollard, C.P. (2004) Do aposematism and

Batesian mimicry require bright colours? A test, using European viper markings. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **271**, 2495-2499.
Xie, G.-x., Ito, E., Maruyama, K., Pietruck, C., Sharma, M., Yu, L.-C. & Pierce Palmer, P. (2000) An alternatively spliced transcript of the rat nociceptin receptor ORL1 gene encodes a truncated receptor. *Molecular Brain Research*, **77**, 1-9.
Xie, H., Wasserman, A., Levine, Z., Novik, A., Grebinskiy, V., Shoshan, A. & Mintz, L. (2002) Large-Scale Protein Annotation through Gene Ontology. *Genome Research*, **12**, 785-794.
Yang, Z. (2015) The BPP program for species tree estimation and species delimitation. *Current Zoology*, **61(5)**, 1-16.
Yeates, D.K., Seago, A., Nelson, L., Cameron, S.L., Joseph, L.E.O. & Trueman, J.W.H. (2011) Integrative taxonomy, or iterative taxonomy? *Systematic Entomology*, **36**, 209-217.
Zapata, F. & Jiménez, I. (2012) Species Delimitation: Inferring Gaps in Morphology across Geography. *Systematic Biology*, **61**, 179-194.
Zeisset, I. & Beebee, T.J.C. (2008) Amphibian phylogeography: a model for understanding historical aspects of species distributions. *Heredity*, **101**, 109-119.
Zhang, C., Rannala, B. & Yang, Z. (2014) Bayesian Species Delimitation Can Be Robust to Guide-Tree Inference Errors. *Systematic Biology*, **63**, 993-1004.
Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**, 2869-2876.
Zheng, Y., Zhao, L., Gao, J. & Fei, Z. (2011) iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequences. *BMC Bioinformatics*, **12**, 1-8.

APPENDIX S2A. Supplementary Tables Chapter 2

Table S2A.1. Mitochondrial and nuclear DNA sequences retrieved from Genbank for gene genealogies and species tree reconstruction.

Species	Loci									# sequences/species (species tree reconstruction)
	12S	16S	SIAH	COI_A	COI_B	CytB	H3	RAG1	tRNA	
O. arborea	1	2	1	1	1	3	1	1	1	**12**
O. granulifera	3	9	1	1	28	30	1	1	3	**77**
O. histrionica	4	6	1	74	3	7	1	1	4	**101**
O. lehmanni	1	1	1	18	1	1	1	1	1	**26**
O. pumilio	6	29	1	21	46	27	0	1	6	**138**
O. speciosa	1	2	1	1	1	4	1	1	1	**13**
O. sylvatica	5	5	1	1	1	3	1	1	5	**23**
O. vicentei	1	1	1	1	1	1	1	1	1	**9**
# sequences/loci	**22**	**55**	**8**	**118**	**82**	**76**	**8**	**8**	**22**	**398**

Accession #	ID	Species	Locus	Type
DQ502036.1	O.arboreaisolate34012S	*O. arborea*	12S	MT
DQ502036.1	O.arboreaisolate340	*O. arborea*	16S	MT
AF128610.1	O.arboreus	*O. arborea*	16S	MT
DQ503060.1	O.arboreusvoucherCWM18636	*O. arborea*	SIAH	NC
AF097504.1	O.arboreus	*O. arborea*	COI_A	MT
DQ502763.1	O.arboreusisolate340	*O. arborea*	COI_B	MT
AF128612.1	O.arboreus1	*O. arborea*	CytB	MT
AF120015.1	O.arboreus2	*O. arborea*	CytB	MT
DQ502467.1	O.arboreusisolate340	*O. arborea*	CytB	MT
DQ502306.1	O.arboreusisolate340	*O. arborea*	H3	NC
DQ503312.1	O.arboreusvoucherCWM18636	*O. arborea*	RAG1	NC
DQ502036.1	O.arboreaisolate340	*O. arborea*	tRNA	MT
DQ502035.1	O.granuliferaisolate33912S	*O. granulifera*	12S	MT
EU342659.1	O.granuliferavoucherKS2012S	*O. granulifera*	12S	MT
EU342658.1	O.granuliferavoucherTNHC64390	*O. granulifera*	12S	MT
DQ502035.1	O.granuliferaisolate339	*O. granulifera*	16S	MT
HE804764.1	O.granuliferamitochondrial1	*O. granulifera*	16S	MT
HE804765.1	O.granuliferamitochondrial2	*O. granulifera*	16S	MT
HE804766.1	O.granuliferamitochondrial3	*O. granulifera*	16S	MT
HE804767.1	O.granuliferamitochondrial4	*O. granulifera*	16S	MT
HE804768.1	O.granuliferamitochondrial5	*O. granulifera*	16S	MT
EU342659.1	O.granuliferavoucherKS20	*O. granulifera*	16S	MT
EU342658.1	O.granuliferavoucherTNHC64390	*O. granulifera*	16S	MT
AF128607.1	O.granuliferus	*O. granulifera*	16S	MT
DQ503059.1	O.granuliferusvoucherCWM19044	*O. granulifera*	SIAH	NC
AF097505.1	O.granuliferus	*O. granulifera*	COI_A	MT

DQ502762.1	O.granuliferusisolate339	*O. granulifera*	COI_B	MT
HQ841105.1	O.granuliferusisolateIW0925	*O. granulifera*	COI_B	MT
HQ841106.1	O.granuliferusisolateIW0927	*O. granulifera*	COI_B	MT
HQ841107.1	O.granuliferusisolateIW0928	*O. granulifera*	COI_B	MT
HQ841108.1	O.granuliferusisolateIW0929	*O. granulifera*	COI_B	MT
HQ841109.1	O.granuliferusisolateIW0930	*O. granulifera*	COI_B	MT
HQ841110.1	O.granuliferusisolateIW0931	*O. granulifera*	COI_B	MT
HQ841111.1	O.granuliferusisolateIW0932	*O. granulifera*	COI_B	MT
HQ841112.1	O.granuliferusisolateIW0933	*O. granulifera*	COI_B	MT
HQ841113.1	O.granuliferusisolateIW0934	*O. granulifera*	COI_B	MT
HQ841114.1	O.granuliferusisolateIW0939	*O. granulifera*	COI_B	MT
HQ841115.1	O.granuliferusisolateIW0940	*O. granulifera*	COI_B	MT
HQ841116.1	O.granuliferusisolateIW0941	*O. granulifera*	COI_B	MT
HQ841117.1	O.granuliferusisolateIW1007	*O. granulifera*	COI_B	MT
HQ841118.1	O.granuliferusisolateIW1009	*O. granulifera*	COI_B	MT
HQ841119.1	O.granuliferusisolateIW1010	*O. granulifera*	COI_B	MT
HQ841120.1	O.granuliferusisolateIW1012	*O. granulifera*	COI_B	MT
HQ841121.1	O.granuliferusisolateIW1014	*O. granulifera*	COI_B	MT
HQ841122.1	O.granuliferusisolateIW1015	*O. granulifera*	COI_B	MT
HQ841123.1	O.granuliferusisolateIW1016	*O. granulifera*	COI_B	MT
HQ841124.1	O.granuliferusisolateIW1017	*O. granulifera*	COI_B	MT
HQ841125.1	O.granuliferusisolateIW1018	*O. granulifera*	COI_B	MT
HQ841126.1	O.granuliferusisolateIW1019	*O. granulifera*	COI_B	MT
HQ841127.1	O.granuliferusisolateIW1020	*O. granulifera*	COI_B	MT
HQ841128.1	O.granuliferusisolateIW1021	*O. granulifera*	COI_B	MT
HQ841129.1	O.granuliferusisolateIW1022	*O. granulifera*	COI_B	MT
HQ841130.1	O.granuliferusisolateIW1024	*O. granulifera*	COI_B	MT
HQ841131.1	O.granuliferusisolateIW1026	*O. granulifera*	COI_B	MT
AF128609.1	O.granuliferus1	*O. granulifera*	CytB	MT
AF120016.1	O.granuliferus2	*O. granulifera*	CytB	MT
DQ502466.1	O.granuliferusisolate339	*O. granulifera*	CytB	MT
HQ841077.1	O.granuliferusisolateIW0925	*O. granulifera*	CytB	MT
HQ841078.1	O.granuliferusisolateIW0927	*O. granulifera*	CytB	MT
HQ841079.1	O.granuliferusisolateIW0928	*O. granulifera*	CytB	MT
HQ841080.1	O.granuliferusisolateIW0929	*O. granulifera*	CytB	MT
HQ841081.1	O.granuliferusisolateIW0930	*O. granulifera*	CytB	MT
HQ841082.1	O.granuliferusisolateIW0931	*O. granulifera*	CytB	MT
HQ841083.1	O.granuliferusisolateIW0932	*O. granulifera*	CytB	MT
HQ841084.1	O.granuliferusisolateIW0933	*O. granulifera*	CytB	MT
HQ841085.1	O.granuliferusisolateIW0934	*O. granulifera*	CytB	MT
HQ841086.1	O.granuliferusisolateIW0939	*O. granulifera*	CytB	MT
HQ841087.1	O.granuliferusisolateIW0940	*O. granulifera*	CytB	MT
HQ841088.1	O.granuliferusisolateIW0941	*O. granulifera*	CytB	MT

HQ841089.1	O.granuliferusisolateIW1007	*O. granulifera*	CytB	MT
HQ841090.1	O.granuliferusisolateIW1009	*O. granulifera*	CytB	MT
HQ841091.1	O.granuliferusisolateIW1010	*O. granulifera*	CytB	MT
HQ841092.1	O.granuliferusisolateIW1012	*O. granulifera*	CytB	MT
HQ841093.1	O.granuliferusisolateIW1014	*O. granulifera*	CytB	MT
HQ841094.1	O.granuliferusisolateIW1015	*O. granulifera*	CytB	MT
HQ841095.1	O.granuliferusisolateIW1016	*O. granulifera*	CytB	MT
HQ841096.1	O.granuliferusisolateIW1017	*O. granulifera*	CytB	MT
HQ841097.1	O.granuliferusisolateIW1018	*O. granulifera*	CytB	MT
HQ841098.1	O.granuliferusisolateIW1019	*O. granulifera*	CytB	MT
HQ841099.1	O.granuliferusisolateIW1020	*O. granulifera*	CytB	MT
HQ841100.1	O.granuliferusisolateIW1021	*O. granulifera*	CytB	MT
HQ841101.1	O.granuliferusisolateIW1022	*O. granulifera*	CytB	MT
HQ841102.1	O.granuliferusisolateIW1024	*O. granulifera*	CytB	MT
HQ841103.1	O.granuliferusisolateIW1026	*O. granulifera*	CytB	MT
DQ502305.1	O.granuliferusisolate339	*O. granulifera*	H3	NC
DQ503311.1	O.granuliferusvoucherCWM19044	*O. granulifera*	RAG1	NC
DQ502035.1	O.granuliferaisolate339	*O. granulifera*	tRNA	MT
EU342659.1	O.granuliferavoucherKS20	*O. granulifera*	tRNA	MT
EU342658.1	O.granuliferavoucherTNHC64390	*O. granulifera*	tRNA	MT
DQ502032.1	O.histrionicaisolate33612S	*O. histrionica*	12S	MT
DQ502105.1	O.histrionicaisolate49812S	*O. histrionica*	12S	MT
EU342662.1	O.histrionicavoucherTNHCFS4879	*O. histrionica*	12S	MT
AF128617.1	O.histrionicus12S	*O. histrionica*	12S	MT
DQ502032.1	O.histrionicaisolate336	*O. histrionica*	16S	MT
DQ502105.1	O.histrionicaisolate498	*O. histrionica*	16S	MT
AF124117.1	O.histrionica	*O. histrionica*	16S	MT
EU342662.1	O.histrionicavoucherTNHCFS4879	*O. histrionica*	16S	MT
AF128616.1	O.histrionicus	*O. histrionica*	16S	MT
HQ290985.1	O.histrionicus1	*O. histrionica*	16S	MT
DQ503083.1	O.histrionicusvoucherMHNUC344	*O. histrionica*	SIAH	NC
KF582727.1	O.histrionicaisolateara1	*O. histrionica*	COI_A	MT
KF582728.1	O.histrionicaisolateara2	*O. histrionica*	COI_A	MT
KF582729.1	O.histrionicaisolateara3	*O. histrionica*	COI_A	MT
KF582730.1	O.histrionicaisolateara4	*O. histrionica*	COI_A	MT
KF582731.1	O.histrionicaisolateara5	*O. histrionica*	COI_A	MT
KF582732.1	O.histrionicaisolateara6	*O. histrionica*	COI_A	MT
KF582721.1	O.histrionicaisolateArus1	*O. histrionica*	COI_A	MT
KF582722.1	O.histrionicaisolateArus2	*O. histrionica*	COI_A	MT
KF582723.1	O.histrionicaisolateArus3	*O. histrionica*	COI_A	MT
KF582724.1	O.histrionicaisolateArus4	*O. histrionica*	COI_A	MT
KF582725.1	O.histrionicaisolateArus5	*O. histrionica*	COI_A	MT
KF582726.1	O.histrionicaisolateArus7	*O. histrionica*	COI_A	MT

KF582712.1	O.histrionicaisolateDelfinaC10Ok	*O. histrionica*	COI_A	MT
KF582710.1	O.histrionicaisolateDelfinaC11	*O. histrionica*	COI_A	MT
KF582705.1	O.histrionicaisolateDelfinaC12	*O. histrionica*	COI_A	MT
KF582706.1	O.histrionicaisolateDelfinaC13	*O. histrionica*	COI_A	MT
KF582702.1	O.histrionicaisolateDelfinaC9	*O. histrionica*	COI_A	MT
KF582709.1	O.histrionicaisolateDelfinaF19	*O. histrionica*	COI_A	MT
KF582703.1	O.histrionicaisolateDelfinaF20	*O. histrionica*	COI_A	MT
KF582719.1	O.histrionicaisolateDelfinaF21	*O. histrionica*	COI_A	MT
KF582715.1	O.histrionicaisolateDelfinaF24	*O. histrionica*	COI_A	MT
KF582713.1	O.histrionicaisolateDelfinaF25	*O. histrionica*	COI_A	MT
KF582714.1	O.histrionicaisolateDelfinaF26	*O. histrionica*	COI_A	MT
KF582704.1	O.histrionicaisolateDelfinaF28	*O. histrionica*	COI_A	MT
KF582720.1	O.histrionicaisolateDelfinaF50	*O. histrionica*	COI_A	MT
KF582708.1	O.histrionicaisolateDelfinaF51	*O. histrionica*	COI_A	MT
KF582711.1	O.histrionicaisolateDelfinaF54	*O. histrionica*	COI_A	MT
KF582716.1	O.histrionicaisolateDelfinaF55	*O. histrionica*	COI_A	MT
KF582707.1	O.histrionicaisolateDelfinaF56	*O. histrionica*	COI_A	MT
KF582718.1	O.histrionicaisolateDelfinaF57	*O. histrionica*	COI_A	MT
KF582717.1	O.histrionicaisolateDelfinaF58	*O. histrionica*	COI_A	MT
KF582752.1	O.histrionicaisolateLehmC87	*O. histrionica*	COI_A	MT
KF582751.1	O.histrionicaisolateLehmC90	*O. histrionica*	COI_A	MT
KF582750.1	O.histrionicaisolateLehmC91	*O. histrionica*	COI_A	MT
KF582749.1	O.histrionicaisolateLehmF79	*O. histrionica*	COI_A	MT
KF582733.1	O.histrionicaisolatePang4	*O. histrionica*	COI_A	MT
KF582734.1	O.histrionicaisolatePang5	*O. histrionica*	COI_A	MT
KF582693.1	O.histrionicaisolatePaya1	*O. histrionica*	COI_A	MT
KF582694.1	O.histrionicaisolatePaya2	*O. histrionica*	COI_A	MT
KF582695.1	O.histrionicaisolatePaya3	*O. histrionica*	COI_A	MT
KF582696.1	O.histrionicaisolatePaya4	*O. histrionica*	COI_A	MT
KF582697.1	O.histrionicaisolatePaya5	*O. histrionica*	COI_A	MT
KF582698.1	O.histrionicaisolatePaya6	*O. histrionica*	COI_A	MT
KF582699.1	O.histrionicaisolatePaya7	*O. histrionica*	COI_A	MT
KF582700.1	O.histrionicaisolatePaya8	*O. histrionica*	COI_A	MT
KF582701.1	O.histrionicaisolatePaya9	*O. histrionica*	COI_A	MT
KF582673.1	O.histrionicaisolatePayasaC59	*O. histrionica*	COI_A	MT
KF582687.1	O.histrionicaisolatePayasaC60	*O. histrionica*	COI_A	MT
KF582688.1	O.histrionicaisolatePayasaC62	*O. histrionica*	COI_A	MT
KF582671.1	O.histrionicaisolatePayasaC63	*O. histrionica*	COI_A	MT
KF582692.1	O.histrionicaisolatePayasaC64	*O. histrionica*	COI_A	MT
KF582685.1	O.histrionicaisolatePayasaC65	*O. histrionica*	COI_A	MT
KF582674.1	O.histrionicaisolatePayasaC66	*O. histrionica*	COI_A	MT
KF582684.1	O.histrionicaisolatePayasaC67	*O. histrionica*	COI_A	MT
KF582669.1	O.histrionicaisolatePayasaC68	*O. histrionica*	COI_A	MT

KF582683.1	O.histrionicaisolatePayasaC69	*O. histrionica*	COI_A	MT
KF582678.1	O.histrionicaisolatePayasaC79	*O. histrionica*	COI_A	MT
KF582666.1	O.histrionicaisolatePayasaC80	*O. histrionica*	COI_A	MT
KF582691.1	O.histrionicaisolatePayasaE44	*O. histrionica*	COI_A	MT
KF582681.1	O.histrionicaisolatePayasaE45	*O. histrionica*	COI_A	MT
KF582670.1	O.histrionicaisolatePayasaE52	*O. histrionica*	COI_A	MT
KF582672.1	O.histrionicaisolatePayasaE72	*O. histrionica*	COI_A	MT
KF582667.1	O.histrionicaisolatePayasaE74	*O. histrionica*	COI_A	MT
KF582668.1	O.histrionicaisolatePayasaE75	*O. histrionica*	COI_A	MT
KF582680.1	O.histrionicaisolatePayasaF69	*O. histrionica*	COI_A	MT
KF582676.1	O.histrionicaisolatePayasaF70	*O. histrionica*	COI_A	MT
KF582675.1	O.histrionicaisolatePayasaF72	*O. histrionica*	COI_A	MT
KF582689.1	O.histrionicaisolatePayasaF73	*O. histrionica*	COI_A	MT
KF582677.1	O.histrionicaisolatePayasaF74	*O. histrionica*	COI_A	MT
KF582686.1	O.histrionicaisolatePayasaF75	*O. histrionica*	COI_A	MT
KF582679.1	O.histrionicaisolatePayasaF76	*O. histrionica*	COI_A	MT
KF582682.1	O.histrionicaisolatePayasaF77	*O. histrionica*	COI_A	MT
KF582690.1	O.histrionicaisolatePayasaF78	*O. histrionica*	COI_A	MT
AF097498.1	O.histrionicus	*O. histrionica*	COI_A	MT
DQ502760.1	O.histrionicusisolate336	*O. histrionica*	COI_B	MT
DQ502816.1	O.histrionicusisolate498	*O. histrionica*	COI_B	MT
HQ841132.1	O.histrionicusisolateIW0845	*O. histrionica*	COI_B	MT
U70154.1	O.histrionicus1	*O. histrionica*	CytB	MT
AF128618.1	O.histrionicus2	*O. histrionica*	CytB	MT
AF120009.1	O.histrionicus3	*O. histrionica*	CytB	MT
HQ290562.1	O.histrionicus4	*O. histrionica*	CytB	MT
DQ502463.1	O.histrionicusisolate336	*O. histrionica*	CytB	MT
DQ502537.1	O.histrionicusisolate498	*O. histrionica*	CytB	MT
HQ841104.1	O.histrionicusisolateIW0845	*O. histrionica*	CytB	MT
DQ502331.1	O.histrionicusisolate498	*O. histrionica*	H3	NC
DQ503332.1	O.histrionicusvoucherMHNUC344	*O. histrionica*	RAG1	NC
DQ502032.1	O.histrionicaisolate336	*O. histrionica*	tRNA	MT
DQ502105.1	O.histrionicaisolate498	*O. histrionica*	tRNA	MT
EU342662.1	O.histrionicavoucherTNHCFS4879	*O. histrionica*	tRNA	MT
HQ290985.1	O.histrionicus	*O. histrionica*	tRNA	MT
DQ502034.1	O.lehmanniisolate33812S	*O. lehmanni*	12S	MT
DQ502034.1	O.lehmanniisolate338	*O. lehmanni*	16S	MT
DQ503058.1	O.lehmannivoucherCWM19050	*O. lehmanni*	SIAH	NC
KF582754.1	O.lehmanniisolateLehmC86	*O. lehmanni*	COI_A	MT
KF582748.1	O.lehmanniisolateLehmC88	*O. lehmanni*	COI_A	MT
KF582747.1	O.lehmanniisolateLehmE77	*O. lehmanni*	COI_A	MT
KF582740.1	O.lehmanniisolateLehmF111	*O. lehmanni*	COI_A	MT
KF582755.1	O.lehmanniisolateLehmF112	*O. lehmanni*	COI_A	MT

KF582756.1	O.lehmanniisolateLehmF113	*O. lehmanni*	COI_A	MT
KF582746.1	O.lehmanniisolateLehmF80	*O. lehmanni*	COI_A	MT
KF582735.1	O.lehmanniisolateRed1	*O. lehmanni*	COI_A	MT
KF582737.1	O.lehmanniisolateRed2	*O. lehmanni*	COI_A	MT
KF582738.1	O.lehmanniisolateRed3	*O. lehmanni*	COI_A	MT
KF582739.1	O.lehmanniisolateRed4_1	*O. lehmanni*	COI_A	MT
KF582736.1	O.lehmanniisolateYello1	*O. lehmanni*	COI_A	MT
KF582741.1	O.lehmanniisolateYello2	*O. lehmanni*	COI_A	MT
KF582753.1	O.lehmanniisolateYello3	*O. lehmanni*	COI_A	MT
KF582742.1	O.lehmanniisolateYello4	*O. lehmanni*	COI_A	MT
KF582743.1	O.lehmanniisolateYello5	*O. lehmanni*	COI_A	MT
KF582744.1	O.lehmanniisolateYello6	*O. lehmanni*	COI_A	MT
KF582745.1	O.lehmanniisolateYello7	*O. lehmanni*	COI_A	MT
DQ502761.1	O.lehmanniisolate338	*O. lehmanni*	COI_B	MT
DQ502465.1	O.lehmanniisolate338	*O. lehmanni*	CytB	MT
DQ502304.1	O.lehmanniisolate338	*O. lehmanni*	H3	NC
DQ503310.1	O.lehmannivoucherCWM19050	*O. lehmanni*	RAG1	NC
DQ502034.1	O.lehmanniisolate338	*O. lehmanni*	tRNA	MT
HQ290988.1	O.pumilio12S	*O. pumilio*	12S	MT
DQ502235.1	O.pumilioisolate131312S	*O. pumilio*	12S	MT
DQ502237.1	O.pumilioisolate131512S	*O. pumilio*	12S	MT
DQ502062.1	O.pumilioisolate36712S	*O. pumilio*	12S	MT
EU342664.1	O.pumiliovoucherMVZ231703	*O. pumilio*	12S	MT
EU342663.1	O.pumiliovoucherTNHCFS4814	*O. pumilio*	12S	MT
AF124120.1	O.pumilio	*O. pumilio*	16S	MT
AF128613.1	O.pumilio1	*O. pumilio*	16S	MT
HQ290988.1	O.pumilio3	*O. pumilio*	16S	MT
DQ768799.1	O.pumiliocolorvariantBastimentos	*O. pumilio*	16S	MT
DQ502235.1	O.pumilioisolate1313	*O. pumilio*	16S	MT
DQ502237.1	O.pumilioisolate1315	*O. pumilio*	16S	MT
DQ502062.1	O.pumilioisolate367	*O. pumilio*	16S	MT
EF597161.1	O.pumiliovoucherCR-BA	*O. pumilio*	16S	MT
EF597162.1	O.pumiliovoucherCR-CA	*O. pumilio*	16S	MT
EF597163.1	O.pumiliovoucherCR-GJ	*O. pumilio*	16S	MT
EF597164.1	O.pumiliovoucherCR-LC	*O. pumilio*	16S	MT
EF597165.1	O.pumiliovoucherCR-PF	*O. pumilio*	16S	MT
EF597167.1	O.pumiliovoucherCR-SI	*O. pumilio*	16S	MT
EF597166.1	O.pumiliovoucherCR-TB	*O. pumilio*	16S	MT
EF597168.1	O.pumiliovoucherCR-TO06	*O. pumilio*	16S	MT
EF597169.1	O.pumiliovoucherCR-UE	*O. pumilio*	16S	MT
EU342664.1	O.pumiliovoucherMVZ231703	*O. pumilio*	16S	MT
EF597170.1	O.pumiliovoucherP-A03	*O. pumilio*	16S	MT
EF597171.1	O.pumiliovoucherP-B38	*O. pumilio*	16S	MT

EF597172.1	O.pumiliovoucherP-C03	*O. pumilio*	16S	MT
EF597173.1	O.pumiliovoucherP-CA01	*O. pumilio*	16S	MT
EF597174.1	O.pumiliovoucherP-E13	*O. pumilio*	16S	MT
EF597175.1	O.pumiliovoucherP-L08	*O. pumilio*	16S	MT
EF597177.1	O.pumiliovoucherP-P05	*O. pumilio*	16S	MT
EF597176.1	O.pumiliovoucherP-PA08	*O. pumilio*	16S	MT
EF597179.1	O.pumiliovoucherP-S02	*O. pumilio*	16S	MT
EF597178.1	O.pumiliovoucherP-SC07	*O. pumilio*	16S	MT
EF597180.1	O.pumiliovoucherP-T01	*O. pumilio*	16S	MT
EU342663.1	O.pumiliovoucherTNHCFS4814	*O. pumilio*	16S	MT
DQ503116.1	O.pumiliovoucherTQM789	*O. pumilio*	SIAH	NC
AF097500.1	O.pumilio	*O. pumilio*	COI_A	MT
EF597181.1	O.pumiliovoucherCR-BA	*O. pumilio*	COI_A	MT
EF597182.1	O.pumiliovoucherCR-CA	*O. pumilio*	COI_A	MT
EF597183.1	O.pumiliovoucherCR-GJ	*O. pumilio*	COI_A	MT
EF597184.1	O.pumiliovoucherCR-LC	*O. pumilio*	COI_A	MT
EF597185.1	O.pumiliovoucherCR-PF	*O. pumilio*	COI_A	MT
EF597187.1	O.pumiliovoucherCR-SI	*O. pumilio*	COI_A	MT
EF597186.1	O.pumiliovoucherCR-TB	*O. pumilio*	COI_A	MT
EF597188.1	O.pumiliovoucherCR-TO06	*O. pumilio*	COI_A	MT
EF597189.1	O.pumiliovoucherCR-UE	*O. pumilio*	COI_A	MT
EF597190.1	O.pumiliovoucherP-A03	*O. pumilio*	COI_A	MT
EF597191.1	O.pumiliovoucherP-B38	*O. pumilio*	COI_A	MT
EF597192.1	O.pumiliovoucherP-C03	*O. pumilio*	COI_A	MT
EF597193.1	O.pumiliovoucherP-CA01	*O. pumilio*	COI_A	MT
EF597194.1	O.pumiliovoucherP-E13	*O. pumilio*	COI_A	MT
EF597195.1	O.pumiliovoucherP-L08	*O. pumilio*	COI_A	MT
EF597197.1	O.pumiliovoucherP-P05	*O. pumilio*	COI_A	MT
EF597196.1	O.pumiliovoucherP-PA08	*O. pumilio*	COI_A	MT
EF597199.1	O.pumiliovoucherP-S02	*O. pumilio*	COI_A	MT
EF597198.1	O.pumiliovoucherP-SC07	*O. pumilio*	COI_A	MT
EF597200.1	O.pumiliovoucherP-T01	*O. pumilio*	COI_A	MT
DQ502907.1	O.pumilioisolate1313	*O. pumilio*	COI_B	MT
DQ502909.1	O.pumilioisolate1315	*O. pumilio*	COI_B	MT
DQ502784.1	O.pumilioisolate367	*O. pumilio*	COI_B	MT
EU934652.1	O.pumilioisolateAg1	*O. pumilio*	COI_B	MT
EU934653.1	O.pumilioisolateAg2	*O. pumilio*	COI_B	MT
EU934654.1	O.pumilioisolateAg3	*O. pumilio*	COI_B	MT
EU934669.1	O.pumilioisolateAl1	*O. pumilio*	COI_B	MT
EU934670.1	O.pumilioisolateAl2	*O. pumilio*	COI_B	MT
EU934641.1	O.pumilioisolateBa1	*O. pumilio*	COI_B	MT
EU934643.1	O.pumilioisolateBa2	*O. pumilio*	COI_B	MT
EU934640.1	O.pumilioisolateBa3	*O. pumilio*	COI_B	MT

EU934642.1	O.pumilioisolateBa4	*O. pumilio*	COI_B	MT
EU934677.1	O.pumilioisolateBn1	*O. pumilio*	COI_B	MT
EU934655.1	O.pumilioisolateCB1	*O. pumilio*	COI_B	MT
EU934656.1	O.pumilioisolateCB2	*O. pumilio*	COI_B	MT
EU934657.1	O.pumilioisolateCB3	*O. pumilio*	COI_B	MT
EU934648.1	O.pumilioisolateCn1	*O. pumilio*	COI_B	MT
EU934662.1	O.pumilioisolateCn2	*O. pumilio*	COI_B	MT
EU934663.1	O.pumilioisolateCn3	*O. pumilio*	COI_B	MT
EU934664.1	O.pumilioisolateCn4	*O. pumilio*	COI_B	MT
EU934676.1	O.pumilioisolateGP3	*O. pumilio*	COI_B	MT
EU934673.1	O.pumilioisolateLS1	*O. pumilio*	COI_B	MT
EU934675.1	O.pumilioisolateLS2	*O. pumilio*	COI_B	MT
EU934674.1	O.pumilioisolateLS3	*O. pumilio*	COI_B	MT
EU934649.1	O.pumilioisolatePa1	*O. pumilio*	COI_B	MT
EU934651.1	O.pumilioisolatePa2	*O. pumilio*	COI_B	MT
EU934650.1	O.pumilioisolatePa3	*O. pumilio*	COI_B	MT
EU934658.1	O.pumilioisolatePo1	*O. pumilio*	COI_B	MT
EU934659.1	O.pumilioisolatePo2	*O. pumilio*	COI_B	MT
EU934660.1	O.pumilioisolatePo3	*O. pumilio*	COI_B	MT
EU934661.1	O.pumilioisolatePo4	*O. pumilio*	COI_B	MT
EU934678.1	O.pumilioisolatePV1	*O. pumilio*	COI_B	MT
EU934679.1	O.pumilioisolatePV2	*O. pumilio*	COI_B	MT
EU934680.1	O.pumilioisolatePV3	*O. pumilio*	COI_B	MT
EU934665.1	O.pumilioisolateSC1	*O. pumilio*	COI_B	MT
EU934667.1	O.pumilioisolateSC2	*O. pumilio*	COI_B	MT
EU934666.1	O.pumilioisolateSC3	*O. pumilio*	COI_B	MT
EU934668.1	O.pumilioisolateSC4	*O. pumilio*	COI_B	MT
EU934644.1	O.pumilioisolateSo1	*O. pumilio*	COI_B	MT
EU934645.1	O.pumilioisolateSo2	*O. pumilio*	COI_B	MT
EU934646.1	O.pumilioisolateSo3	*O. pumilio*	COI_B	MT
EU934647.1	O.pumilioisolateSo4	*O. pumilio*	COI_B	MT
EU934681.1	O.pumilioisolateSq1	*O. pumilio*	COI_B	MT
EU934682.1	O.pumilioisolateTo1	*O. pumilio*	COI_B	MT
EU934671.1	O.pumilioisolateUy3	*O. pumilio*	COI_B	MT
EU934672.1	O.pumilioisolateUy4	*O. pumilio*	COI_B	MT
U70141.1	O.pumilio1	*O. pumilio*	CytB	MT
U70150.1	O.pumilio10	*O. pumilio*	CytB	MT
U70151.1	O.pumilio11	*O. pumilio*	CytB	MT
U70152.1	O.pumilio12	*O. pumilio*	CytB	MT
AF120011.1	O.pumilio13	*O. pumilio*	CytB	MT
AF128615.1	O.pumilio14	*O. pumilio*	CytB	MT
HQ290565.1	O.pumilio15	*O. pumilio*	CytB	MT
U70142.1	O.pumilio2	*O. pumilio*	CytB	MT

U70143.1	O.pumilio3	*O. pumilio*	CytB	MT
U70144.1	O.pumilio4	*O. pumilio*	CytB	MT
U70145.1	O.pumilio5	*O. pumilio*	CytB	MT
U70146.1	O.pumilio6	*O. pumilio*	CytB	MT
U70147.1	O.pumilio7	*O. pumilio*	CytB	MT
U70148.1	O.pumilio8	*O. pumilio*	CytB	MT
U70149.1	O.pumilio9	*O. pumilio*	CytB	MT
DQ502668.1	O.pumilioisolate1313	*O. pumilio*	CytB	MT
DQ502493.1	O.pumilioisolate367	*O. pumilio*	CytB	MT
EF597201.1	O.pumiliovoucherCR-BA	*O. pumilio*	CytB	MT
EF597202.1	O.pumiliovoucherCR-CA	*O. pumilio*	CytB	MT
EF597203.1	O.pumiliovoucherCR-GJ	*O. pumilio*	CytB	MT
EF597204.1	O.pumiliovoucherCR-LC	*O. pumilio*	CytB	MT
EF597205.1	O.pumiliovoucherCR-PF	*O. pumilio*	CytB	MT
EF597207.1	O.pumiliovoucherCR-SI	*O. pumilio*	CytB	MT
EF597206.1	O.pumiliovoucherCR-TB	*O. pumilio*	CytB	MT
EF597208.1	O.pumiliovoucherCR-TO06	*O. pumilio*	CytB	MT
EF597209.1	O.pumiliovoucherCR-UE	*O. pumilio*	CytB	MT
EF597210.1	O.pumiliovoucherP-A03	*O. pumilio*	CytB	MT
EU325918.1	O.pumiliovoucherTTR117	*O. pumilio*	RAG1	NC
HQ290988.1	O.pumilio	*O. pumilio*	tRNA	MT
DQ502235.1	O.pumilioisolate1313	*O. pumilio*	tRNA	MT
DQ502237.1	O.pumilioisolate1315	*O. pumilio*	tRNA	MT
DQ502062.1	O.pumilioisolate367	*O. pumilio*	tRNA	MT
EU342664.1	O.pumiliovoucherMVZ231703	*O. pumilio*	tRNA	MT
EU342663.1	O.pumiliovoucherTNHCFS4814	*O. pumilio*	tRNA	MT
DQ502037.1	O.speciosaisolate34112S	*O. speciosa*	12S	MT
DQ502037.1	O.speciosaisolate341	*O. speciosa*	16S	MT
AF128595.1	O.speciosus	*O. speciosa*	16S	MT
DQ503061.1	O.speciosusvoucherCWM17826(D)	*O. speciosa*	SIAH	NC
AF097503.1	O.speciosus	*O. speciosa*	COI_A	MT
AF097503.1	O.speciosus	*O. speciosa*	COI_B	MT
AF128597.1	O.speciosus1	*O. speciosa*	CytB	MT
AF120014.1	O.speciosus2	*O. speciosa*	CytB	MT
DQ502468.1	O.speciosus3	*O. speciosa*	CytB	MT
U70153.1	O.speciosusisolate341	*O. speciosa*	CytB	MT
DQ502307.1	O.speciosusisolate341	*O. speciosa*	H3	NC
DQ503313.1	O.speciosusvoucherCWM17826(D)	*O. speciosa*	RAG1	NC
DQ502037.1	O.speciosaisolate341	*O. speciosa*	tRNA	MT
DQ502059.1	O.sylvaticaisolate36412S	*O. sylvatica*	12S	MT
EU342661.1	O.sylvaticavoucherQCAZ16562	*O. sylvatica*	12S	MT
EU342660.1	O.sylvaticavoucherQCAZ27423	*O. sylvatica*	12S	MT
HQ290990.1	O.sylvaticus12S	*O. sylvatica*	12S	MT

AY364569.1	O.sylvaticusvoucherQCAZ16563	*O. sylvatica*	12S	MT
DQ502059.1	O.sylvaticaisolate364	*O. sylvatica*	16S	MT
EU342661.1	O.sylvaticavoucherQCAZ16562	*O. sylvatica*	16S	MT
EU342660.1	O.sylvaticavoucherQCAZ27423	*O. sylvatica*	16S	MT
HQ290990.1	O.sylvaticus	*O. sylvatica*	16S	MT
AY364569.1	O.sylvaticusvoucherQCAZ16563	*O. sylvatica*	16S	MT
DQ503065.1	O.sylvaticusvoucherLSUMZ14730	*O. sylvatica*	SIAH	NC
DQ502781.1	O.sylvaticusisolate364	*O. sylvatica*	COI_A	MT
DQ502781.1	O.sylvaticusisolate364	*O. sylvatica*	COI_B	MT
HQ290567.1	O.sylvaticus1	*O. sylvatica*	CytB	MT
AF324041.1	O.sylvaticus2	*O. sylvatica*	CytB	MT
DQ502490.1	O.sylvaticusisolate364	*O. sylvatica*	CytB	MT
DQ502312.1	O.sylvaticusisolate364	*O. sylvatica*	H3	NC
DQ503317.1	O.sylvaticusvoucherLSUMZ14730	*O. sylvatica*	RAG1	NC
DQ502059.1	O.sylvaticaisolate364	*O. sylvatica*	tRNA	MT
EU342661.1	O.sylvaticavoucherQCAZ16562	*O. sylvatica*	tRNA	MT
EU342660.1	O.sylvaticavoucherQCAZ27423	*O. sylvatica*	tRNA	MT
HQ290990.1	O.sylvaticus	*O. sylvatica*	tRNA	MT
AY364569.1	O.sylvaticusvoucherQCAZ16563	*O. sylvatica*	tRNA	MT
DQ502167.1	O.vicenteiisolate114812S	*O. vicentei*	12S	MT
DQ502167.1	O.vicenteiisolate1148	*O. vicentei*	16S	MT
DQ503115.1	O.vicenteivoucherKRL789	*O. vicentei*	SIAH	NC
DQ502869.1	O.vicenteiisolate1148	*O. vicentei*	COI_A	MT
DQ502869.1	O.vicenteiisolate1148	*O. vicentei*	COI_B	MT
DQ502602.1	O.vicenteiisolate1148	*O. vicentei*	CytB	MT
DQ502365.1	O.vicenteiisolate1148	*O. vicentei*	H3	NC
DQ503364.1	O.vicenteivoucherKRL789	*O. vicentei*	RAG1	NC
DQ502167.1	O.vicenteiisolate1148	*O. vicentei*	tRNA	MT

Table S2A.2. Estimates of DNA sequence diversity, nucleotide evolution models and neutrality test results. For further details, see the Methodology section in manuscript.

Partition	Total number of sites*	Polymorphic sites (S)	k	Nucleotide diversity (Pi)	Nucleotide evolution model	Tajima's D**	Codon Partition Scheme
12S	2374	306	76.307	0.032	GTR+I+G	-0.674	Step 2 (1-2-3)
16S	481	63	12.114	0.025	TIM2+I+G	-0.722	Step 2 (1-2-3)
SIAH1	395	4	1.143	0.003	JC	-1.434	Step 2 (1-2-3)
COI(a)	416	89	15.748	0.038	HKY+G	-0.467	Step 1 (1-2)(3)
COI(b)	593	147	41.756	0.070	TrN+I+G	0.901	Start scheme (1)(2)(3)
CytB	249	79	13.823	0.056	TPM2uf+G	-0.775	Start scheme (1)(2)(3)
H3	303	7	3.000	0.010	TrN+I	-0.427	Step 1 (1-2)(3)

Rag1	432	7	2.190	0.005	TPM2uf	-1.208	Step 1 (1-2)(3)
tRNA-val	2388	367	87.268	0.037	TIM2+I+G	-0.856	Step 2 (1-2-3)

k: Average number of nucleotide differences
*Excluding sites with gaps / missing data
** All p vales were >0.10.

Table S2A.3. Species occurrence points used for ENMs reconstructions and multivariate niche analysis. I refrain of publicly provide geographic locations for *O. histrionica* and *O. lehmanni* due high levels of wildlife trafficking and negative impact in natural populations (Lynch & Arroyo, 2009). Further information for academic purposes could be requested to the authors.

Species	Longitude	Latitude
O. arborea	-82.1676	8.9102
O. arborea	-82.2010	8.7852
O. arborea	-82.2426	8.8186
O. arborea	-82.2260	8.8852
O. arborea	-82.2093	8.8436
O. arborea	-82.1676	8.8352
O. arborea	-82.1760	8.9436
O. arborea	-82.2260	8.7686
O. arborea	-82.2010	8.9436
O. arborea	-82.2010	8.8519
O. arborea	-82.2093	8.9436
O. arborea	-82.2176	8.7602
O. arborea	-82.1760	8.8269
O. arborea	-82.1593	8.8602
O. arborea	-82.2426	8.8602
O. arborea	-82.2176	8.9436
O. arborea	-82.2093	8.7519
O. arborea	-82.1926	8.9269
O. arborea	-82.2093	8.8602
O. arborea	-82.2093	8.9352
O. arborea	-82.2093	8.8269
O. arborea	-82.2260	8.7602
O. arborea	-82.2176	8.7769
O. arborea	-82.2093	8.9269
O. arborea	-82.2176	8.7436
O. arborea	-82.2426	8.8852
O. arborea	-82.2176	8.8686
O. arborea	-82.2510	8.8102

O. arborea	-82.2426	8.8436
O. arborea	-82.2510	8.9102
O. arborea	-82.1676	8.8602
O. arborea	-82.2176	8.7602
O. arborea	-82.1926	8.7936
O. arborea	-82.2010	8.8602
O. arborea	-82.2426	8.7769
O. arborea	-82.2343	8.9102
O. arborea	-82.2343	8.9186
O. arborea	-82.1926	8.8852
O. arborea	-82.2426	8.8936
O. arborea	-82.2176	8.9519
O. granulifera	-84.2333	9.6125
O. granulifera	-84.1518	9.4562
O. granulifera	-84.1189	9.5811
O. granulifera	-83.8898	9.2537
O. granulifera	-83.7000	9.3833
O. granulifera	-83.5833	8.4833
O. granulifera	-83.5259	8.7145
O. granulifera	-83.5191	8.6959
O. granulifera	-83.5092	8.6975
O. granulifera	-83.5369	8.6333
O. granulifera	-83.4927	8.7120
O. granulifera	-83.4886	8.7005
O. granulifera	-83.4861	8.6916
O. granulifera	-83.4861	8.6993
O. granulifera	-83.4833	8.7897
O. granulifera	-83.3333	8.9667
O. granulifera	-83.2545	8.8987
O. granulifera	-83.9642	9.3839
O. granulifera	-83.8442	9.3119
O. granulifera	-83.8022	9.3119
O. granulifera	-83.7123	9.2339
O. granulifera	-83.4963	8.8620
O. granulifera	-83.3763	8.8260
O. granulifera	-83.9342	9.4079
O. granulifera	-83.5683	8.9460
O. granulifera	-83.5097	8.7153
O. granulifera	-83.5007	8.6879
O. granulifera	-83.4471	8.8098
O. pumilio	-84.6225	10.4762

O. pumilio	-84.5806	10.4284
O. pumilio	-84.5646	10.3833
O. pumilio	-84.5500	10.3833
O. pumilio	-84.4944	10.3614
O. pumilio	-84.4833	10.3667
O. pumilio	-84.4833	10.3667
O. pumilio	-84.4799	10.3614
O. pumilio	-84.4749	10.3745
O. pumilio	-84.4725	10.3614
O. pumilio	-84.4667	10.4667
O. pumilio	-84.1500	10.4167
O. pumilio	-84.1011	10.4321
O. pumilio	-84.0821	10.2600
O. pumilio	-84.0361	10.4547
O. pumilio	-84.0313	10.5463
O. pumilio	-84.0291	10.4099
O. pumilio	-84.0209	10.4358
O. pumilio	-84.0200	10.4244
O. pumilio	-84.0167	10.5571
O. pumilio	-84.0167	10.4667
O. pumilio	-84.0167	10.4531
O. pumilio	-84.0167	10.4667
O. pumilio	-84.0133	10.4261
O. pumilio	-84.0128	10.4467
O. pumilio	-84.0068	10.4304
O. pumilio	-84.0060	10.4505
O. pumilio	-84.0058	10.4308
O. pumilio	-84.0044	10.3262
O. pumilio	-84.0036	10.4312
O. pumilio	-83.9833	10.4333
O. pumilio	-83.9715	10.5115
O. pumilio	-83.8831	10.7664
O. pumilio	-83.7833	10.4500
O. pumilio	-83.7737	10.2147
O. pumilio	-83.7628	10.7736
O. pumilio	-83.7486	10.2106
O. pumilio	-83.7413	10.2108
O. pumilio	-83.6349	10.3637
O. pumilio	-83.5965	9.9223
O. pumilio	-83.5863	9.9293
O. pumilio	-83.5830	10.1670

O. pumilio	-83.5733	10.4192
O. pumilio	-83.5500	9.8167
O. pumilio	-83.5316	10.5867
O. pumilio	-83.5298	10.5865
O. pumilio	-83.5279	10.5865
O. pumilio	-83.5199	9.7803
O. pumilio	-83.5197	10.5250
O. pumilio	-83.5167	10.5724
O. pumilio	-83.5167	10.5833
O. pumilio	-83.5052	10.0993
O. pumilio	-83.4941	10.0947
O. pumilio	-83.4906	10.0893
O. pumilio	-83.4905	10.0893
O. pumilio	-83.4833	10.0833
O. pumilio	-83.4776	10.4712
O. pumilio	-83.4283	10.3846
O. pumilio	-83.3987	10.0924
O. pumilio	-83.3985	10.0924
O. pumilio	-83.2833	10.0333
O. pumilio	-83.0925	9.8995
O. pumilio	-83.0500	9.9833
O. pumilio	-83.0493	9.4348
O. pumilio	-83.0416	9.4511
O. pumilio	-83.0255	9.4656
O. pumilio	-83.0236	9.6722
O. pumilio	-83.0236	9.6722
O. pumilio	-83.0031	9.6542
O. pumilio	-82.9667	9.7333
O. pumilio	-82.9667	9.7333
O. pumilio	-82.9385	9.5718
O. pumilio	-82.9333	9.5667
O. pumilio	-82.8552	9.5604
O. pumilio	-82.8266	9.7156
O. pumilio	-82.4977	9.2250
O. pumilio	-82.3227	9.4178
O. pumilio	-82.2479	9.3616
O. pumilio	-82.2416	9.3607
O. pumilio	-82.2117	9.3622
O. pumilio	-82.1668	9.3453
O. pumilio	-82.1500	9.3167
O. pumilio	-81.9171	8.9774

O. pumilio	-81.8865	9.1819
O. pumilio	-84.7756	10.5462
O. pumilio	-84.7456	10.9121
O. pumilio	-84.1757	10.8641
O. pumilio	-83.7677	10.6541
O. pumilio	-83.8757	11.0561
O. pumilio	-84.4756	11.2060
O. speciosa	-82.2926	8.9102
O. speciosa	-82.3260	8.7602
O. speciosa	-82.3260	8.8852
O. speciosa	-82.2510	8.7352
O. speciosa	-82.2760	8.8519
O. speciosa	-82.3010	8.9019
O. speciosa	-82.2593	8.7519
O. speciosa	-82.3510	8.8519
O. speciosa	-82.2926	8.7352
O. speciosa	-82.2760	8.8186
O. speciosa	-82.2176	8.7436
O. speciosa	-82.3676	8.9102
O. speciosa	-82.3676	8.9436
O. speciosa	-82.2426	8.7352
O. speciosa	-82.3010	8.7519
O. speciosa	-82.2176	8.8269
O. speciosa	-82.2843	8.7686
O. speciosa	-82.3426	8.8519
O. speciosa	-82.2760	8.7436
O. speciosa	-82.3426	8.9686
O. speciosa	-82.2593	8.7769
O. speciosa	-82.2093	8.7269
O. speciosa	-82.2426	8.8102
O. speciosa	-82.2260	8.7519
O. speciosa	-82.3593	8.9269
O. speciosa	-82.3176	8.7436
O. speciosa	-82.2843	8.8519
O. speciosa	-82.3260	8.8936
O. speciosa	-82.2260	8.8102
O. speciosa	-82.3093	8.7186
O. speciosa	-82.2426	8.8436
O. speciosa	-82.3176	8.9436
O. speciosa	-82.2760	8.8436
O. speciosa	-82.3176	8.9102

O. speciosa	-82.2676	8.8102
O. speciosa	-82.2843	8.8602
O. speciosa	-82.2843	8.8769
O. speciosa	-82.2426	8.7686
O. speciosa	-82.3093	8.9352
O. speciosa	-82.3426	8.9352
O. sylvatica	-79.7161	0.3440
O. sylvatica	-79.7157	0.3442
O. sylvatica	-79.7094	0.3512
O. sylvatica	-79.7090	0.3516
O. sylvatica	-79.6335	0.3836
O. sylvatica	-79.6151	0.3842
O. sylvatica	-79.5386	0.6617
O. sylvatica	-79.4673	0.3319
O. sylvatica	-79.3810	-0.0104
O. sylvatica	-79.3702	-0.3874
O. sylvatica	-79.3653	-0.5861
O. sylvatica	-79.3627	-0.5883
O. sylvatica	-79.3592	-0.0565
O. sylvatica	-79.3529	-0.4987
O. sylvatica	-79.3334	-0.5662
O. sylvatica	-79.3261	1.0033
O. sylvatica	-79.2451	-0.2387
O. sylvatica	-79.2311	-0.2377
O. sylvatica	-79.2029	0.4775
O. sylvatica	-79.2029	0.4770
O. sylvatica	-79.2007	0.4766
O. sylvatica	-79.2003	0.4766
O. sylvatica	-79.1998	0.4764
O. sylvatica	-79.1992	0.4761
O. sylvatica	-79.1990	0.4766
O. sylvatica	-79.1978	0.4761
O. sylvatica	-79.1964	0.4770
O. sylvatica	-79.1635	-0.2778
O. sylvatica	-79.1355	0.1054
O. sylvatica	-79.0486	0.0966
O. sylvatica	-78.9980	0.9084
O. sylvatica	-78.9463	-0.1063
O. sylvatica	-78.9137	1.0678
O. sylvatica	-78.8943	1.3525
O. sylvatica	-78.7737	1.1334

O. sylvatica	-78.7599	1.2335
O. sylvatica	-78.7560	1.1832
O. sylvatica	-78.6906	1.0874
O. sylvatica	-78.6789	1.0767
O. sylvatica	-78.6534	1.0727
O. sylvatica	-78.6488	1.0720
O. sylvatica	-78.6443	1.0646
O. sylvatica	-78.6298	1.0612
O. sylvatica	-78.6250	1.0007
O. sylvatica	-78.6244	1.0427
O. sylvatica	-78.6182	1.0474
O. sylvatica	-78.6138	1.0323
O. sylvatica	-78.6116	1.0317
O. sylvatica	-78.6098	1.0332
O. sylvatica	-78.5998	0.8668
O. sylvatica	-78.5941	1.0636
O. sylvatica	-78.5832	0.9169
O. sylvatica	-78.5808	0.9126
O. sylvatica	-78.5662	0.9171
O. sylvatica	-78.5091	0.9010
O. vicentei	-80.6000	8.6667
O. vicentei	-80.6114	8.6661
O. vicentei	-80.5979	8.6386
O. vicentei	-80.6228	8.6585
O. vicentei	-80.5625	8.6696
O. vicentei	-80.6007	8.6979
O. vicentei	-80.5524	8.6951
O. vicentei	-80.6271	8.6980
O. vicentei	-80.6153	8.6737
O. vicentei	-80.6004	8.6305
O. vicentei	-80.5989	8.6144
O. vicentei	-80.6477	8.6276
O. vicentei	-80.5961	8.6238
O. vicentei	-80.5784	8.6322
O. vicentei	-80.6373	8.6307
O. vicentei	-80.5523	8.6996
O. vicentei	-80.6429	8.6468
O. vicentei	-80.6094	8.7061
O. vicentei	-80.6383	8.7033
O. vicentei	-80.5607	8.6574

Table S2A.4. Bioclimatic variables used for ENMs and multivariate niche analysis (Hijmans *et al.*, 2005).

ID	Description*	Selected
BIO1	Annual mean temperature	Yes
BIO2	Mean diurnal range (max temp – min temp) (monthly average)	Yes
BIO3	Isothermality (BIO1/BIO7) * 100	Yes
BIO4	Temperature Seasonality (Coefficient of Variation)	Yes
BIO5	Max Temperature of Warmest Period	No
BIO6	Min Temperature of Coldest Period	No
BIO7	Temperature Annual Range (BIO5-BIO6)	No
BIO8	Mean Temperature of Wettest Quarter	No
BIO9	Mean Temperature of Driest Quarter	No
BIO10	Mean Temperature of Warmest Quarter	No
BIO11	Mean Temperature of Coldest Quarter	No
BIO12	Annual Precipitation	Yes
BIO13	Precipitation of Wettest Period	No
BIO14	Precipitation of Driest Period	No
BIO15	Precipitation Seasonality (Coefficient of Variation)	Yes
BIO16	Precipitation of Wettest Quarter	No
BIO17	Precipitation of Driest Quarter	No
BIO18	Precipitation of Warmest Quarter	Yes
BIO19	Precipitation of Coldest Quarter	No
Altitude	SRTM 90m Digital Elevation Data	Yes

*Temperature data is given in units of °C*10 (with 0.1 °C precision). Precipitation data is in mm.

Table S2A.5. Values of niche filling as estimated on G- and E- space for 8 *Oophaga* species. Details are presented in the methods.

	Geographic (G-space)			Environmental (E-space)***		
Species	**# grids predicted***	**# grids geographic range****	**Niche Filling**	**Predicted niche**	**Known occupied niche**	**Niche Filling**
O. arborea	9137	1189	13.01%	496096	58770	11.85%
O. granulifera	7157	6861	95.86%	318014	291769	91.75%
O. histrionica	24503	24503	100.00%	127073	127073	100.00%
O. lehmanni	6141	756	12.31%	216496	62825	29.02%
O. pumilio	80577	31944	39.64%	448449	399455	89.07%
O. speciosa	4963	1726	34.78%	288446	262180	90.89%
O. sylvatica	26207	24109	91.99%	198025	168664	85.17%
O. vicentei	8027	2843	35.42%	343769	208902	60.77%

**Binary suitability scores from Maxent niche models using as a threshold the sensitivity-specificity sum maximization approach (Liu et al., 2005).*

*** Estimated using a broad species range as previously described in methodology*

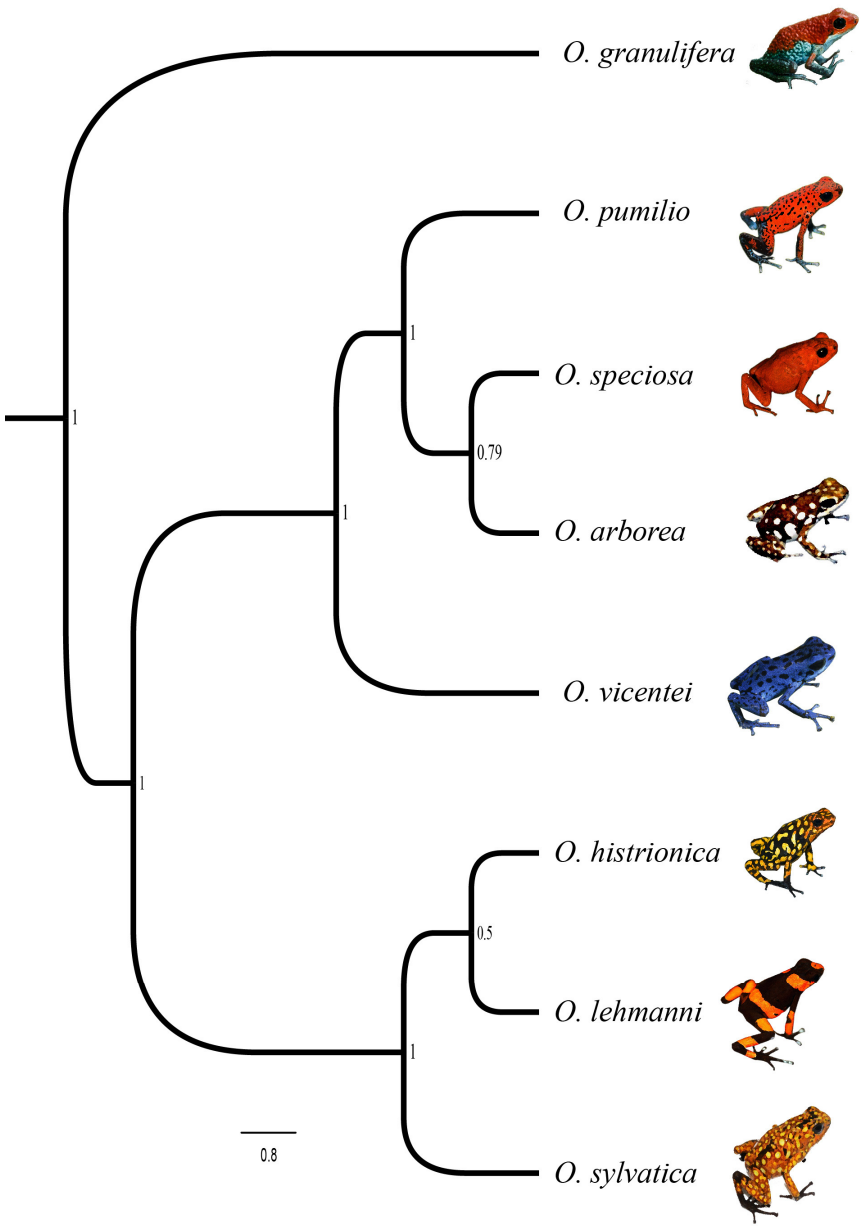
****Area of convex hulls (pixels) enclosing predicted and occupied species observations after PCA.*

CITES

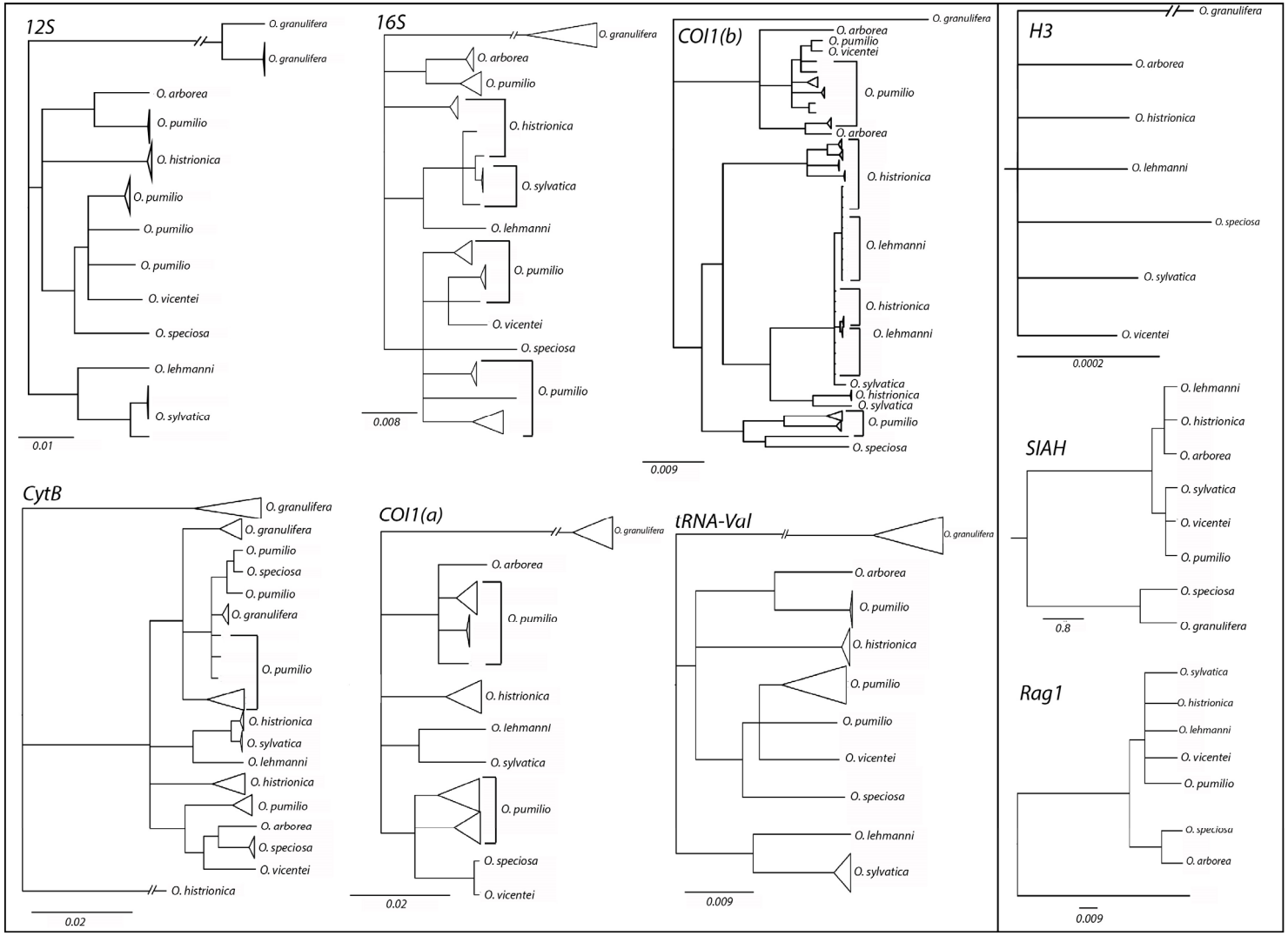
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978.

Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385-393.

Lynch, J.D. & Arroyo, S.B. (2009) Risks to Colombian Amphibian Fauna from Cultivation of Coca (*Erythroxylum coca*): A Geographical Analysis. *Journal of Toxicology and Environmental Health, Part A*, **72**, 974-985.



Supplementary Fig. S2B.1. Reconstructed Bayesian species tree considering alternative partition codon schemes (i.e PARTITIONFINDER, see Appendix S2C for details). Numbers on each node represent Bayesian posterior probability (BPP) support.



Supplementary Fig. S2B.2. Mitochondrial (left) and nuclear (right) gene genealogies as inferred by ML and Bayesian methods (same topology). Tree branches are collapsed to 70% bootstrap support or 0.7 posterior probability (PP). Size of triangles at the end of the tips are proportional to the number of clustered individuals in that particular tip.

APPENDIX S2C: Supplementary Methods Chapter 2

Phylogenetic analyses: Gene genealogies and species tree estimation

For each locus I aligned all sequences using the Clustal-W algorithm (Thompson *et al.*, 1994). The resulting alignments were used to estimate the model of nucleotide substitution using the sample size-corrected Akaike Information Criterion (AICc) (Posada & Buckley, 2004) as implemented in JMODELTEST v2.1.4 (Posada, 2008). The best partitioning scheme for each loci was identified with PartitionFinder 1.1.1 (Lanfear *et al.*, 2012) using the greedy algorithm and the BIC criterion (Table S2A.2). Gene genealogies (Appendix S2B Fig. S2B.2) were constructed using maximum likelihood (ML) and Bayesian inference (BI) as follows: first, using MEGA6 (Tamura *et al.*, 2013) and the results derived from JMODELTEST (Table S1.2), I obtained ML genealogies for each gene using extensive (level 5) Subtree Pruning and Regrafting (SPR) heuristic searches. The relative support for each node was estimated by generating 1,000 bootstrap replicates. Second, I constructed BI-based genealogies using BEAST v. 2.1.2 (Drummond *et al.*, 2012; Bouckaert *et al.*, 2014) with a relaxed molecular clock and an uncorrelated log-normal (UCLN) model of molecular rate heterogeneity. For each gene, I ran three chains for 10 million generations sampled every 100 steps. The resulting trees and log files of the three independent runs were combined using LOGCOMBINER 2.1.2. and resampled at a lower frequency (i.e. 400 steps or ¼) to reduce the number of final trees to 52.500 (Bouckaert *et al.*, 2014). For each estimated parameter, convergence was assessed using TRACER 1.6 (http://tree.bio.ed.ac.uk/software/tracer/) and effective samples sizes (ESS) were calculated to ensure adequate mixing (ESS>300, after 30% burn in). I summarized the posterior probability density of the combined tree and log files as a maximum clade credibility (MCC) tree using TREEANNOTATOR 2.1.2 (Bouckaert *et al.*, 2014). I visualized all trees using FIGTREE 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

Single gene genealogies revealed the lack of species monophyly and the existence of genealogical conflict among loci. Therefore, I reconstructed the species phylogeny using two methods that account for coalescence stochasticity: *BEAST (Drummond & Rambaut, 2007; Drummond *et al.*, 2012; Bouckaert *et al.*, 2014) and the Minimize Deep Coalescence (MDC) method (Maddison & Knowles, 2006), as implemented in PHYLONET 5.3.2 (Than & Nakhleh, 2009). Neither of these methods account for possible gene flow between populations or species. However, migration rates less 0.1 have little effect on species delimitation (Eckert & Carstens, 2008; Yang, 2015) .While MDC finds the species tree that minimizes the number of extra lineages across loci, *BEAST is a Bayesian method that uses MCMC to account for uncertainty in both the individual gene and species tree. Both methods assume that the sampled loci are neutral and that there is no recombination within them. To test these assumptions, I estimated Tajima's D (Tajima, 1989) using DNASP 5.10.1 (Librado & Rozas, 2009) and tested for intralocus recombination using the difference of sums of squares method (DSS) as implemented in TOPALI 2 (Milne *et al.*, 2009). I implemented * BEAST analyses using a relaxed clock by using an uncorrelated lognormal model of molecular evolutionary rate heterogeneity. Because gene genealogies were inferred using different datasets (*i.e.* individuals), I considered each locus as a partition. The * BEAST species tree was estimated using a uniform Yules Process prior and a piecewise constant population size model (gamma, shape= 1, scale= 5). Three separate chains were run for 300 million iterations sampling

parameters every 10,000 generations. Independent runs and resulting trees and log files were combined and visualized as described above and 30% as burn-in. All Bayesian analyses were run in the Bugaboo Dell Xeon cluster of the western Canada's WestGrid computer facilities (www.westgrid.ca). Pairwise species genetic distances were estimated as the unweighted mean *p*-distance (uncorrected nucleotide difference) for all sample loci.

Bioclimatic niche modelling

The amount of reliable presence data of *Oophaga* frogs is relatively limited. Thus, I modelled species' niches (ENM) using maximum entropy (MAXENT 3.3.3k) (Phillips *et al.*, 2006), a method that performs relatively well with small sample sizes (Hernandez *et al.*, 2006; Pearson *et al.*, 2007). To minimize over-fitting problems due to collinearity, I conducted PCA- and uncorrelated (UC-) analyses (Peterson & Nakazawa, 2008; Ahmadzadeh *et al.*, 2013). First, all bioclimatic layers were cropped to include the known ranges of all species using ARCGIS 10.1 (ESRI). I conducted a principal components analysis (PCA) on the correlation matrix of the cropped, standardized variables (19 bioclimatic variables and altitude) and used the principal components (PCs) with eigenvalues greater than one (Kaiser, 1960) (n= 4, Table S3.6) as predictors in subsequent ENMs. Second, I removed all highly related bioclimatic variables (R > 0.90, n= 12), and used the remaining 8 non-correlated climatic layers (Altitude, Bio1, Bio2, Bio3, Bio4, Bio12, Bio18; Hijmans *et al.*, 2005) as predictors in subsequent ENMs (Table S1.4). Because PCA- and UC- based models were statistically identical (D>0.85, p<0.05 for all 8 species)(Warren *et al.*, 2010) I focused on latter ones. This allowed me to evaluate the relative contribution of each climatic variable to the ENMs and to improve interpretability.

To further minimize over-fitting I also performed species-specific tuning (Merow *et al.*, 2013; Radosavljevic & Anderson, 2014) by comparing models (n_{total}= 332) with pseudo-absences (n=10,000) taken from different backgrounds defined by the union area of circles of radius of 50, 100, 150, 200 and 300 km around each occurrence point (VanDerWal *et al.*, 2009). Models were run under the "*autofeatures"* option in MAXENT (Phillips *et al.*, 2006) with regularization multipliers from 0.5 to 4.0 in increments of 0.5. Model selection was based on the corrected Akaike Information Criterion which has shown to outperform other methods (Warren *et al.*, 2010; Warren & Seifert, 2011; Hijmans, 2012) (AICc; Table S3.7). For each species, I ran 10 replicates of the selected model and used the average across replicates for my downstream analyses. In every case the selected models showed very good predicting ability (area under the ROC curve, AUC-training > 0.92, AUC-testing > 0.92). All tuning analyses were performed using ENMTOOLS 1.4.3 (Warren *et al.*, 2010).

Geographic correlates of diversification: range overlap and niche filling

I estimated age-range correlations (ARC) (Berlocher & Feder, 2002; Fitzpatrick & Turelli, 2006) between genetic distance and geographic overlap (Warren's D) (Warren *et al.*, 2008) as implemented in Fitzpatrick and Turrelli (2006) and the R package 'phyloClim' (Heibl & Calenge, 2013). The geographic overlap between each pair of species was estimated as the area of the intersect polygons between two species (divided by the area of the species with the smallest distribution (Nakazato *et al.*, 2010). In addition, to examine the relationship between genetic distance and geographic overlap, I used Principal Coordinates of Neighbour Matrices

(PCNM) as implemented in 'vegan' R package (Oksanen *et al.*, 2013). I retained all set of eigenfunctions with positive Moran's I coefficients and assessed the significance of the *r* statistic using 1000 random permutations of the response data (Legendre *et al.*, 2011).

For each species, I calculated the degree of "niche filling" (the proportion of the space predicted to suit the species' ecological requirements that is actually known to be occupied by each species) by comparing the geographical extension of the estimated "suitable habitat" (suitability pixels) to the known geographical range inferred from the empirical occurrence records (G-space). In order to maximize the agreement between observed and predicted distributions, "suitable habitat" threshold was defined based on the sensitivity-specificity sum maximization approach, which has been shown to perform well for this purpose (Liu *et al.*, 2005). Additionally, I estimated the degree of niche filling in environmental space (E-space) (Soberón & Peterson, 2011) by PCA of the 8 non-correlated bioclimatic variables at the predicted suitability grids (i.e by Maxent) and the currently known species distributions. Niche filling was accessed by comparing the area of convex hulls enclosing predicted suitability and currently known distribution observations in the first two multivariate dimensions (PCA1 and PCA2) which explained >85% of total variation in every case.

Niche differentiation *vs.* conservatism: sister-species comparisons

First, for each pair of species I calculated niche overlap using the Schoener's D metrics in the ENMTOOLS package (Warren *et al.*, 2010). Then, I used a background test (Warren *et al.*, 2008) to determine whether the estimated D values ranging from zero (no overlap) to one (total overlap) were more similar than expected by chance. Because overlap measurements vary depending on which species is considered to be the focal one, this process is repeated in the opposite direction, generating two null distributions per analysis. A potential caveat of this test is that the definition of geographic range can bias the likelihood of finding niche conservatism (or divergence) by altering the relative amount of potentially unsuitable habitats included in the range of a taxon. Although addressing this issue is beyond the scope of this paper, I assessed the robustness of the niche overlap estimates by comparing the results of background tests using "broad" and "narrow" species ranges. Based on the relative low dispersal capability of *Oophaga* frogs (Nowakowski *et al.*, 2013), I defined these ranges as the geographic area covered by a convex Hull polygon including either the occurring localities (narrow range) or a 5 Km radius around each occurrence point (broad range).

Secondly, because ENM estimates of niche divergence might overlook small ecological differences, I also implemented a multivariate method (McCormack *et al.*, 2010). For each of the environmental variables, I drew data from both presence points and 1,000 random background points from within the background zones using DIVA-GIS 7.5 (http://www.diva-gis.org/; (Hijmans *et al.*, 2001). Then, I used PCA of the correlation matrix to reduce the number of variables to four independent niche axes (PCs), each of them explaining at least 4% of the total variance. For each pair of species, on each single axis (PC) I tested for patterns of niche divergence (or conservatism) by comparing the observed mean difference between the PC-scores to a bootstrap distribution (n=1,000) of mean differences in background values (McCormack *et al.*, 2010). I considered that niche divergence was supported only if both the observed niche divergence itself and the background tests were significant (McCormack *et al.*,

2010; Wooten *et al.*, 2013).

Table S2C.6. PCA descriptions of the 20 bioclimatic variables (Table S1.4) used to reduce the number of variables for ENM's modelling and analysis.

PC	r Eigen Value	Percent of EigenValues	Accumulative of EigenValues
1	1171622.00	62.72	62.72
2	476560.60	25.51	88.23
3	177937.10	9.53	97.76
4	25698.08	1.38	99.13
5	10721.04	0.57	99.70
6	3895.72	0.21	99.91
7	747.60	0.04	99.95
8	395.86	0.02	99.97
9	218.81	0.01	99.99
10	115.59	0.01	99.99
11	69.51	0.00	100.00
12	34.01	0.00	100.00
13	13.62	0.00	100.00
14	10.35	0.00	100.00
15	8.36	0.00	100.00
16	3.07	0.00	100.00
17	1.20	0.00	100.00
18	0.38	0.00	100.00
19	0.14	0.00	100.00
20	0.00	0.00	100.00

Table S2C.7. ENMs parameters of the selected models based on their predicting ability.

Species	R*	Background	AICc score	AUC Training	AUC Testing
O. arborea	0.5	150 km	926.890	0.998	0.998
O. granulifera	1	300 km	576.511	0.990	0.980
O. histrionica	0.5	300 km	2179.788	0.984	0.982
O. lehmanni	1	300 km	840.433	0.992	0.993
O. pumilio	1	300 km	1971.144	0.957	0.956
O. speciosa	0.5	300 km	794.101	0.997	0.997
O. sylvatica	1	300 km	1184.471	0.977	0.978
O. vicentei	0.5	200 km	896.849	0.997	0.991

*Regularization parameter used by Maxent.

CITES

Ahmadzadeh, F., Flecks, M., Carretero, M.A., Böhme, W., Ilgaz, C., Engler, J.O., James Harris, D., Üzüm, N. & Rödder, D. (2013) Rapid lizard radiation lacking niche conservatism: ecological diversification within a complex landscape. *Journal of Biogeography*, **40**, 1807-1818.

Berlocher, S.H. & Feder, J.L. (2002) Sympatric speciation in phytophagous insects: Moving Beyond Controversy? *Annual Review of Entomology*, **47**, 773-815.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, **10**, e1003537.

Drummond, A. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology*, **7**, 214.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-1973.

Eckert, A.J. & Carstens, B.C. (2008) Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Molecular Phylogenetics and Evolution*, **49**, 832-842.

Fitzpatrick, B.M. & Turelli, M. (2006) The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution*, **60**, 601-615.

Heibl, C. & Calenge, C. (2013) Integrating phylogenetics and climatic niche modeling. In: *http://cran.r-project.org/web/packages/phyloclim/phyloclim.pdf*

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773-785.

Hijmans, R., Guarino, L., Cruz, M. & Rojas, E. (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*, 15-19.

Hijmans, R.J. (2012) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, **93**, 679-688.

Lanfear, R., Calcott, B., Ho, S.Y.W. & Guindon, S. (2012) PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution*, **29**, 1695-1701.

Legendre, P., Oksanen, J. & ter Braak, C.J.F. (2011) Testing the significance of canonical axes in redundancy analysis. *Methods in Ecology and Evolution*, **2**, 269-277.

Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451-1452.

Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385-393.

Maddison, W.P. & Knowles, L.L. (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, **55**, 21-30.

McCormack, J.E., Zellmer, A.J. & Knowles, L.L. (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: insights from tests with niche models. *Evolution*, **64**, 1231-1244.

Merow, C., Smith, M.J. & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, **36**, 1058-1069.

Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D.F. & Wright, F. (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, **25**, 126-127.

Nakazato, T., Warren, D.L. & Moyle, L.C. (2010) Ecological and geographic modes of species divergence in wild tomatoes. *American Journal of Botany*, **97**, 680-693.

Nowakowski, A.J., Otero Jiménez, B., Allen, M., Diaz-Escobar, M. & Donnelly, M.A. (2013) Landscape resistance to movement of the poison frog, *Oophaga pumilio*, in the lowlands of northeastern Costa Rica. *Animal Conservation*, **16**, 188-197.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P. & Minchin, P. (2013) vegan: Community Ecology Package. R package version 2.0-6. *Available at: http://mirror.bjtu.edu.cn/cran/web/packages/vegan/*,

Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Townsend Peterson, A. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102-117.

Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with Solenopsis invicta and Solenopsis richteri. *Global Ecology and Biogeography*, **17**, 135-144.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.

Posada, D. (2008) jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, **25**, 1253-1256.

Posada, D. & Buckley, T.R. (2004) Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. *Systematic Biology*, **53**, 793-808.

Radosavljevic, A. & Anderson, R.P. (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography*, **41**, 629-643.

Soberón, J. & Peterson, A.T. (2011) Ecological niche shifts and environmental space anisotropy: a cautionary note. *Revista Mexicana de Biodiversidad*, **82**, 1384-1355.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.
Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, **30**, 2725-2729.
Than, C. & Nakhleh, L. (2009) Species Tree Inference by Minimizing Deep Coalescences. *PLoS Comput Biol*, **5**, e1000501.
Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.
VanDerWal, J., Shoo, L.P., Graham, C. & Williams, S.E. (2009) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, **220**, 589-594.
Warren, D.L. & Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335-342.
Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883.
Warren, D.L., Glor, R.E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607-611.
Wooten, J.A., Camp, C.D., Combs, J.R., Dulka, E., Reist, A. & Walker, D.M. (2013) Re-evaluating niche conservatism versus divergence in the Woodland Salamander genus Plethodon: a case study of the parapatric members of the Plethodon glutinosus species complex. *Canadian Journal of Zoology*, **91**, 883-892.
Yang, Z. (2015) The BPP program for species tree estimation and species delimitation. *Current Zoology*, **61(5)**, 1-16.

APPENDIX S3A: Supplementary Tables Chapter 3

Table S3A.1. Description of 49 morphologic variables used for the estimation of morphological clusters and combinatory species delimitation.

Type	Position	Variable	Description of measurement	Units/Type
Morphometric	Dorsal/ventral	**Snout-cloaca length** (av_lenght)	Mean value of dorsal/ventral distance between the tip of the nose and the cloaca.	mm
Morphometric	Dorsal	**Inter-orbital width** (eye)	Width of the skull between the ocular ridges	mm
Morphometric	Dorsal	**Dorsal area** (d_area)	Dorsal area without including limbs	mm^2
Morphometric	Dorsal	**Head Depth** (snout)	depth of the head at its deepest point immediately posterior to the orbital cavities	mm
Coloration pattern	Dorsal	**Number of coloured dorsal spots** (nspots_d)	Count of spots with bright colouration in dorsal surface	count
Coloration pattern	Dorsal	**Median dorsal colour spot area** (d_spots_median)	Median value for the area of dorsal spots/individual	mm^2
Coloration pattern	Dorsal	**Average of dorsal colour spot area** (d_spots_mean)	Mean value for the area of dorsal spots/individual	mm^2
Coloration pattern	Dorsal	**Standard deviation of dorsal colour spot area** (d_spots_std)	Standard deviation value for the area of dorsal spots/individual	mm^2
Coloration pattern	Dorsal	**Percentage of dorsal colouration** (%_col_d)	Summatory of individual dorsal colour spot areas divided by the total dorsal area	%
Coloration pattern	Dorsal	**Dorsal coloured spots longitudinal centrality** (xmean_c)	Mean distance from inter-pupilar Y axes to the centroid of each coloured spot in the X axes.	mm
Coloration pattern	Dorsal	**Standard deviation longitudinal centrality** (xcenter_std)	Standard deviation value for Dorsal coloured spots longitudinal centrality	mm
Coloration pattern	Dorsal	**Dorsal coloured spots transversal centrality** (ymean_c)	Mean distance from inter-pupilar Y axes to the centroid of each coloured spot in the Y axes.	mm
Coloration pattern	Dorsal	**Standard deviation transversal centrality** (ycenter_std)	Standard deviation value for Dorsal coloured spots transversal centrality	mm
Coloration pattern	Dorsal	**Mean dorsal circularity** (cd_mean)	Mean value for dorsal coloured spots circularity ($4*\pi*Area/Perimeter^2$)	no units
Coloration pattern	Dorsal	**Standard deviation dorsal spots circularity** (circ_d_std)	Standard deviation value for dorsal coloured spots circularity	no units
Coloration pattern	Dorsal	**Mean dorsal roundness** (rd_mean)	Mean value for dorsal coloured spots roundness ($4*\pi*Area/Feret^2$)**	no units
Coloration pattern	Dorsal	**Standard deviation dorsal spots roundness** (roun_d_std)	Standard deviation value for dorsal coloured spots roundness	no units
Coloration pattern	Ventral	**Number of coloured ventral spots** (nspots_v)	Count of spots with bright colouration in ventral surface	count
Coloration pattern	Ventral	**Ventral area** (v_area)	Ventral area without including limbs	mm^2
Coloration pattern	Ventral	**Median ventral colour spot area** (v_spots_median)	Median value for the area of ventral spots/individual	mm^2

Coloration pattern	Ventral	**Average of ventral colour spot area** (v_spots_mean)	Mean value for the area of ventral spots/individual	mm^2
Coloration pattern	Ventral	**Standard deviation of ventral colour spot area** (v_spots_std)	Standard deviation value for the area of ventral spots/individual	mm^2
Coloration pattern	Ventral	**Percentage of ventral colouration** (%_col_v)	Summatory of individual ventral colour spot areas divided by the total dorsal area	%
Coloration pattern	Ventral	**Ventral coloured spots longitudinal centrality** (xvmean_c)	Mean distance from inter-pupilar Y axes to the centroid of each coloured spot in the X axes (Ventral)	mm
Coloration pattern	Ventral	**Standard deviation ventral longitudinal centrality** (xvcenter_std)	Standard deviation value for Dorsal coloured spots longitudinal centrality (Ventral)	mm
Coloration pattern	Ventral	**Ventral coloured spots transversal centrality** (yvmean_c)	Mean distance from inter-pupilar Y axes to the centroid of each coloured spot in the Y axes (Ventral)	mm
Coloration pattern	Ventral	**Standard deviation ventral transversal centrality** (yvcenter_std)	Standard deviation value for ventral coloured spots transversal centrality (Ventral)	mm
Coloration pattern	Ventral	**Mean ventral circularity** (cv_mean)	Mean value for ventral coloured spots circularity ($4*\pi*Area/Perimeter^2$)	no units
Coloration pattern	Ventral	**Standard deviation dorsal spots circularity** (circ_v_std)	Standard deviation value for ventral coloured spots circularity	no units
Coloration pattern	Ventral	**Mean ventral roundness** (rv_mean)	Mean value for dorsal coloured spots roundness ($4*\pi*Area/Feret^2$)**	no units
Coloration pattern	Ventral	**Standard deviation dorsal spots roundness** (round_v_std)	Standard deviation value for ventral coloured spots roundness	no units
Coloration intensity	Dorsal	**Dorsal Background red intensity** (red_b_d)	R value (RGB component) dorsal background colouration	RGB Units
Coloration intensity	Dorsal	**Dorsal Background green intensity** (green_b_d)	G value (RGB component) dorsal background colouration	RGB Units
Coloration intensity	Dorsal	**Dorsal Background blue intensity** (blue_b_d)	B value (RGB component) dorsal background colouration	RGB Units
Coloration intensity	Dorsal	**Dorsal Background Brightness Index** (indexbrigh_ind_b_d)	Dorsal background colouration (brightness index calculated as 0.299R+0.587G+0.114B)	RGB Units
Coloration intensity	Dorsal	**Dorsal coloured red intensity** (red_d_d)	R value (RGB component) dorsal coloured colouration	RGB Units
Coloration intensity	Dorsal	**Dorsal coloured green intensity** (green_d_d)	G value (RGB component) dorsal coloured colouration	RGB Units
Coloration intensity	Dorsal	**Dorsal coloured blye intensity** (blue_d_d)	B value (RGB component) dorsal coloured colouration	RGB Units
Coloration intensity	Dorsal	**Dorsal coloured Brightness Index** (brigh_ind_d_d)	Dorsal bright colouration (brightness index calculated as 0.299R+0.587G+0.114B))	RGB Units
Coloration intensity	Dorsal	**Dorsal Brightness difference** (brigh_diff_d)	Dorsal Brightness difference (brigh_ind_d_d) - (brigh_ind_b_d)	RGB Units
Coloration intensity	Ventral	**Ventral Background red intensity** (red_b_d)	R value (RGB component) ventral background colouration	RGB Units

Coloration intensity	Ventral	**Ventral Background green intensity** (green_b_d)	G value (RGB component) ventral background colouration	RGB Units
Coloration intensity	Ventral	**Ventral Background blue intensity** (blue_b_d)	B value (RGB component) ventral background colouration	RGB Units
Coloration intensity	Ventral	**Ventral Background Brightness Index** (indexbrigh_ind_b_d)	Ventral background colouration (brightness index)	RGB Units
Coloration intensity	Ventral	**Ventral coloured red intensity** (red_d_d)	R value (RGB component) ventral coloured colouration	RGB Units
Coloration intensity	Ventral	**Ventral coloured green intensity** (green_d_d)	G value (RGB component) ventral coloured colouration	RGB Units
Coloration intensity	Ventral	**Ventral coloured blye intensity** (blue_d_d)	B value (RGB component) ventral coloured colouration	RGB Units
Coloration intensity	Ventral	**Ventral coloured Brightness Index** (brigh_ind_d_d)	Ventral coloured colouration (brightness index)	RGB Units
Coloration intensity	Ventral	**Ventral Brightness difference** (brigh_diff_d)	Ventral Brightness difference (brigh_ind_d_d) - (brigh_ind_b_d)	RGB Units

Table S3A.2. Description of nuclear markers implemented in this study.

Loci	Forward Sequence (5'-3')	Reverse Sequence (5'-3')	Type	PCR amplicon	Estimated Size
4_hist	TAGCAATGTCGCTGGATGAC	ACAAAGCACCGCACTACAT	microsatellite	Yes	156-170
13_hist	TCCAGTTTGTGAGGTAGCAAAA	AAAACGGGAAAGATTTATTCTGA	microsatellite	Yes	160-190
44_hist	TCACTGCTTCATTAATTGTT	GAATCCAGCAGCCTCTGTAAG	microsatellite	Yes	196-290
63_hist	CCCAGGGTAGAGTTGCTTTTT	TTCGCTTTTCTTCCCTTTCA	microsatellite	Yes	190-240
88_hist	CGGTCCACCTCTTTCCTGT	ATGCAAGCTTCCTGCTCTGT	microsatellite	No	205-233
147_hist	ACCTTATACATCTGCACGCTTTC	TCATTATCCAACGACATGCTTCC	microsatellite	No	240-285
109_hist	TGTGTTTGTATGCGGGTACG	TGCCATGGTGTGTGATGAG	microsatellite	No	400-418
110_hist	TCACCTTTGCATCCAGTTCA	TGGTATTGTGCTGTTCCACTTT	microsatellite	Yes	198-260
1696_hist	GGACTATCCGCAATTTGCAT	CCATTTTGAGGGAGTGAGGA	microsatellite	No	160-190
3024_hist	TGTGAGAAATATTTGGAGGCTGG	GTCTGAATACTTTCCGTACCCAC	microsatellite	No	150-198
5133_hist	ATTAAAGCAATATAAGCCGGCCC	TCCGCAATTTGTCATGGTTCTAG	microsatellite	No	280-330
6868_hist	ACGGGCTAAGACTGCGATATAG	CACGCTATGGGTTCAACTG	microsatellite	Yes	165-190
comp324	GCTCTTCCGATCTCTGTGTT	TGTTTTCCAGAAACAAACCA	tbn-DNA	Yes	400-500
comp328	TGAACTCACCAGTAAACCTTTG	AATAGCCCTTGGTTGTATGG	tbn-DNA	No	400-500
comp332	CAAGTGCAGTCATGAAAACC	CCTCACATGTCCTGTGTCTC	tbn-DNA	Yes	400-500
comp344	TCACAAGACCTCCAGCAGTA	AGACAGCAGTTCCAGTGTGA	tbn-DNA	No	400-500
comp350	AAGCACTTCAAACCACAAGAC	AGCCAGAGGATGTGTAAACG	tbn-DNA	Yes	400-500
comp365	GGGGAGGAGAGAAAAATCAT	CACATTTCCCATTTCGTACA	tbn-DNA	No	400-500
comp378	TGTAGTTTTCCTATATTTCCTTCAC	TAGTGAATGGATCCCTCTGG	tbn-DNA	Yes	400-500
comp380	GTGAACAGTCATTTTGTGCTG	CCCATCCACTACAGATCACA	tbn-DNA	No	400-500
comp396	ACAGGGCCTAATGCTGAG	CTGACTGCACTTTGGTTCTG	tbn-DNA	Yes	400-500
comp406	GTGGTGTTCTTGGGAAGC	ATTTGGTTCAATTGTGTTGC	tbn-DNA	No	400-500
comp425	GGAATGATCTTTTGGTTCGT	GCCTCCCATCTAAACAGAAT	tbn-DNA	Yes	400-500
comp430	TCGTTCTCAAGAAACCCTGT	AGCATAGCAACCGAGTACCT	tbn-DNA	No	400-500
comp438	CATCCTTCTCTTCAGGCACT	TAAAGGCCATTTTAGGGACA	tbn-DNA	No	400-500
comp_462	AAAACAAGCTCCTGTCTGCT	AGAAGCTTTGGTTCCATCTG	tbn-DNA	No	400-500
comp444	TCATCCATCAGCTTAACTCCT	TCCATTCCATTAGATTGATCTTC	tbn-DNA	Yes	400-500
comp480	TTGGAACATAGTCATTCGTCA	CACGTAAAACGGTAGGCTGT	tbn-DNA	Yes	400-500
comp484	TGACCTTATATAGGCAGTGCAG	TCACCCAGATTCCTTTATCC	tbn-DNA	No	400-500
comp493	GGGTGCAGTGATCTGTGAG	TCAACCACAAGAAGAACCTC	tbn-DNA	Yes	400-500
comp511	CAGCCTTACACAGGGATTTG	CAACATCAAGAAATCGGCTA	tbn-DNA	Yes	400-500

comp534	CATCTTATCTTACCCAAATTTTCC	GTGCTCTTCCGATCTGTTGT	tbn-DNA	No	400-500
comp539	TCAAGATTCACTGGTGCAAG	CCTCTTCTGGACATGGAGAT	tbn-DNA	No	400-500
comp540	TATGTAAGAATGAGGTGGAGACA	CGTTCACCTATATATGCACACTT	tbn-DNA	No	400-500
comp547	AATTACCAGCCAGACTCCAG	GATGTTAAGCGTCACAGCAT	tbn-DNA	No	400-500
comp548	AGGCGTATAAGTGGATTTGC	CAGGAGCTGGTCTTGTAACC	tbn-DNA	No	400-500
comp553	ACTTTGATGCAGTCTCACACA	CACATAGCAGTGCTTCAAAAA	tbn-DNA	No	400-500
comp562	GGACTCTACCTTGGGCAGTA	GTGCAAAAGTGTCACCAAAA	tbn-DNA	Yes	400-500
comp583	CGAAGAGACACCTCCCATTA	CGAGAGCTGCACTCACTTC	tbn-DNA	Yes	400-500
comp594	GTAAACCCGCAGGTGAACT	TAAAGTGACACGCTTTGGTG	tbn-DNA	Yes	400-500

Table S3A.3. Quality descriptors of the assembled transcriptome used for nuclear loci markers design.

Species	*O. lehmanni*
File size (Mb)	77.3
# sequences (Contigs)	117,082
Total bases (b)	66,367,430
Mean sequence length (b)	566.8
Minimum length (b)	201
Maximum length (b)	12,906
N50 contig size (b)	809
N75 contig size (b)	353

APPENDIX S3B: Supplementary Methods Chapter 3

Mitochondrial genetic clustering

To determine the number of mitochondrial clusters (K_{MT}), phylogenetic trees were constructed under different Bayesian (BI) and maximum likelihood (ML) optimality criteria. First, using MEGA6 (Tamura et al. 2013), I obtained an ML genealogy using extensive (level 5) Subtree Pruning and Regrafting (SPR) heuristic searches. The relative support for each node was estimated by generating 1,000 bootstrap replicates. Second, I constructed a BI-based tree using BEAST v. 2.1.2 (Drummond et al. 2012; Bouckaert et al. 2014) with a relaxed molecular clock and an uncorrelated log-normal (UCLN) model of molecular rate heterogeneity. I ran three chains for 10 million generations sampled every 100 steps. The resulting trees and log files of the three independent runs were combined using LOGCOMBINER 2.1.2. and resampled at a lower frequency (i.e. 400 steps or ¼) to reduce the number of final trees to 52.500 (Bouckaert et al. 2014). For each estimated parameter, convergence was assessed using TRACER 1.6 (http://tree.bio.ed.ac.uk/software/tracer/) and effective samples sizes (ESS) were calculated to ensure adequate mixing (ESS>320, after 30% burn in). I summarized the posterior probability density of the combined tree and log files as a maximum clade credibility (MCC) tree using TREEANNOTATOR 2.1.2 (Bouckaert et al. 2014). I visualized all trees using FIGTREE 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/). The phylogenetic relationship among sampling locations was plotted onto a TCS haplotype network with a connection limit of 95% (Clement et al. 2000) implemented in the program *PopART* (Population Analysis with reticulate trees) available at http://popart.otago.ac.nz/. Finally, to delimit independent evolving linages I performed a Bayesian implementation of the PTP model for species delimitation (Zhang et al. 2013) at the bTPT online server (http://species.h-its.org/) with MCMC generations = 1,000,000, thinning = 1000 and Burn-in = 0.1.

Microsatellite molecular markers

Microsatellite markers were generated as follow: genomic DNA from a single individual of the known *O. histrionica* nominal species was isolated as described in the Methods section of the main manuscript. The genomic sample was sequenced on one-eighth of a plate using 454GS-FLX technology (Roche Applied Science) at the Cornell University's BioResource Center. I obtained 14,992 fragment reads with lengths ranging from 150 to 825 bases. I searched for di-, tri- and tetra-nucleotide tandem repeats from the dataset using MSATCOMMANDER (Faircloth 2008; Pino et al. 2012). The online version of PRIMER3 (Rozen and Skaletsky 2000) was used to access the suitability for annealing temperature and single primer-pair for each loci were chosen based on the highest PRIMER3 assigned score. Primer pairs were tested using conventional PCR in twelve individuals from two geographic localities and those loci that were successfully amplified (Table S3A.3 in Appendix S3A) were used to genotype 310 individuals from 15 different geographic locations (Table S3A.1 in Appendix S3A). I followed previously described protocols for multiplex PCR conditions, allele detection and allele calling (Anttila et al. 2014).

Nuclear loci for multiplex PCR

First, I generated a skin transcriptome by Trizol (Invitrogen™) extraction of total RNA from skin portions of one known individual of *O. lehmanni*. Paired-end libraries (150-mer, x2; insert lengths ~300 bp) were synthesized using the Genomic Sample Prep kit (Illumina™) according to manufacturer's instructions. Library quality was assessed using a 2100 Bioanalyzer (Agilent). Libraries were sequenced on an Illumina Hi-Seq 2000 with a paired-end module. Recognition, sorting, trimming of tags and quality control of the raw DNA-sequence reads were performed in FLEXBAR (Dodt et al. 2012). The remaining reads (~128 million) and the TRINITY platform (Haas et al. 2013) were used for the generation of a *de – novo* transcriptome. Further information about transcriptome composition and quality is presented in Table S3A.4 in Appendix S3A provided in Supplementary Material .

Allele-frequency genetic clustering

In order to assign individuals of the *O. histrionica* complex to genetic clusters, I implemented Bayesian clustering algorithms using a combination of allele frequency data from microsatellite (n=6) and transcriptome-derived haplotypes (n=13) in STRUCTURE v. 2.3.4 (Pritchard et al. 2000; Hubisz et al. 2009). A model assuming admixture and correlated allele frequencies was implemented in 10 independent runs with one million MCMC iterations (burn-in = 100000) for a series of clusters (K_{NUC}) ranging from 1 (panmixia) to 15 (maximum number of localities sampled). Then, I used the statistic ΔK to select the value of K with the uppermost hierarchical level of population structure in the data (Evanno et al. 2005). For detecting species clusters and individual assignment while treating K_{NUC} as a random variable, I also implemented STRUCTURAMA v. 1.0 (Huelsenbeck et al. 2011). The models were run for 10 million generations, sampling every 100th cycle and a burn-in period of 10000 samples. A different set of prior values was used for the shape and scale of the gamma distribution of K_{NUC} (shape: scale = 2.5:0.5, 2.5:1, 5:2, 10:1) (Edwards and Knowles 2014). I chose the K_{NUC} with the highest probability and individuals were assigned to species cluster by using the posterior probability of cluster assignment. Finally, for the multivariate method the combined allele frequency data was converted to synthetic variables by PCoA of a Nei's genetic distance matrix as implemented in GenAlex v. 6.5 (Peakall and Smouse 2012) and clusters were visually inspected from PCoA dispersion plots.

CITES

Anttila K., Couturier C.S., Øverli Ø., Johnsen A., Marthinsen G., Nilsson G.E. , Farrell A.P. 2014. Atlantic salmon show capability for cardiac acclimation to warm temperatures. Nat Commun. 5

Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A. , Drummond A.J. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput. Biol. 10: e1003537.

Clement M., Posada D. , Crandall K.A. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9: 1657-1659.

Dodt M., Roehr J., Ahmed R. , Dieterich C. 2012. FLEXBAR—Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. Biology. 1: 895-905.

Drummond A.J., Suchard M.A., Xie D. , Rambaut A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29: 1969-1973.

Edwards D.L. , Knowles L.L. 2014. Species detection and individual assignment in species delimitation: can integrative data increase efficacy? Proc. R. Soc. Lond., Ser. B: Biol. Sci. 281: 20132765.
Evanno G., Regnaut S. , Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. 14: 2611-2620.
Faircloth B.C. 2008. msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol. Ecol. Resour. 8: 92-94.
Haas B.J., Papanicolaou A., Yassour M., Grabherr M., Blood P.D., Bowden J., Couger M.B., Eccles D., Li B., Lieber M., MacManes M.D., Ott M., Orvis J., Pochet N., Strozzi F., Weeks N., Westerman R., William T., Dewey C.N., Henschel R., LeDuc R.D., Friedman N. , Regev A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protocols. 8: 1494-1512.
Hubisz M.J., Falush D., Stephens M. , Pritchard J.K. 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9: 1322-1332.
Huelsenbeck J.P., Andolfatto P. , Huelsenbeck E.T. 2011. Structurama: Bayesian Inference of Population Structure. Evol Bioinform Online. 7: 55-59.
Peakall R. , Smouse P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 28: 2537-2539.
Pino J., Ascunce M., Reed D. , Hugot J.-P. 2012. Development and characterization of microsatellite markers for the endangered Laotian rock-rat (Laonastes aenigmamus) using 454-sequencing technology. Conservation Genetics Resources. 4: 999-1002.
Pritchard J.K., Stephens M. , Donnelly P. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics. 155: 945-959.
Rozen S. , Skaletsky H. 2000. Primer3 on the World Wide Web for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totow, NJ.
Tamura K., Stecher G., Peterson D., Filipski A. , Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30: 2725-2729.
Zhang J., Kapli P., Pavlidis P. , Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics. 29: 2869-2876.

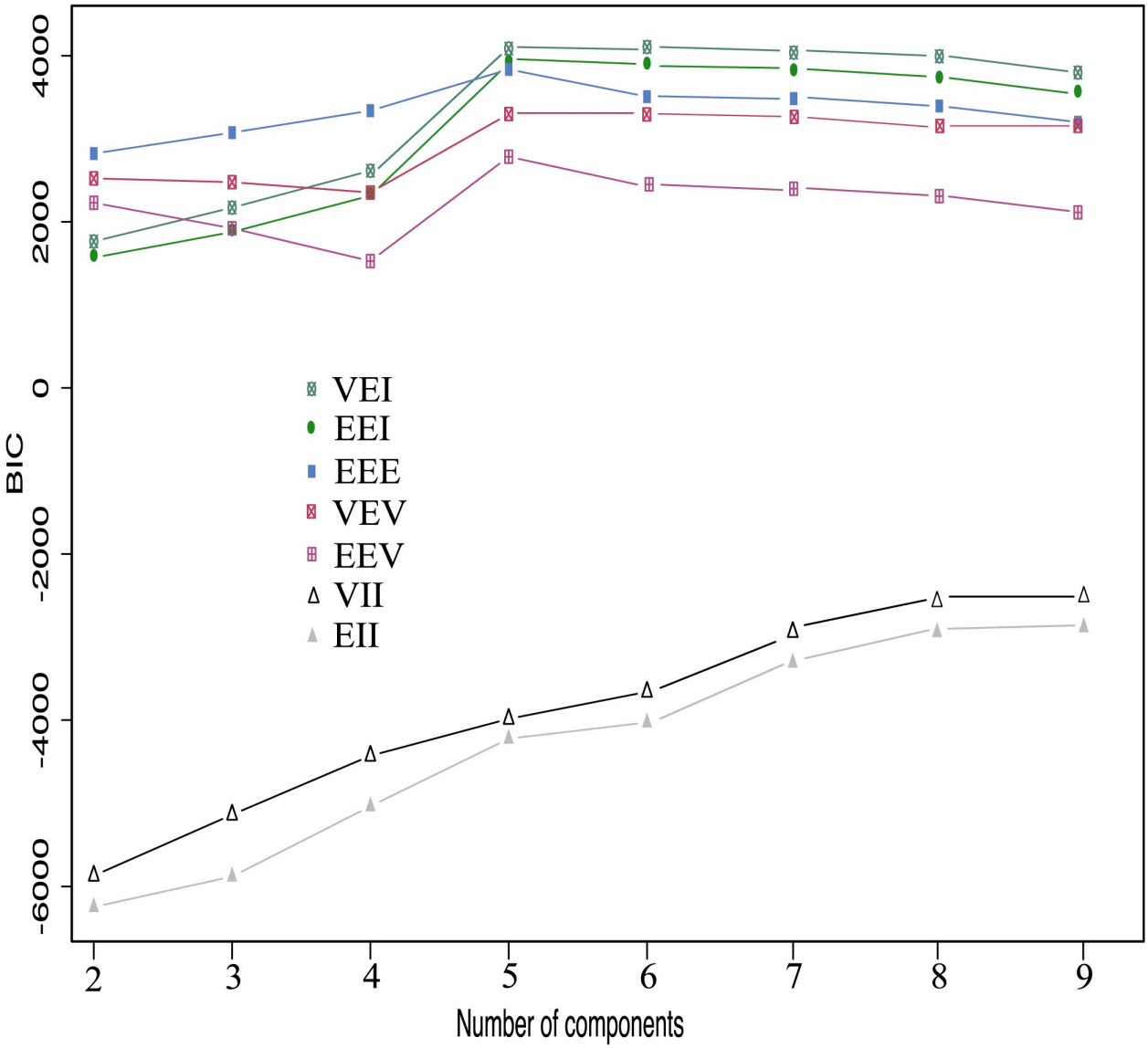


Figure S3C.6. Gaussian clustering graph showing a plateau in BIC value at K=5. According to the *Mclust* analysis, the best model corresponded to an equal- covariance and diagonal distribution model (VEI) with five components or clusters.

APPENDIX S4A: Supplementary Tables Chapter 4

Table S4A.1 Number of reads and type of library (pared vs. single-end) for each individual RNA-seq experiment.

Individual library	Raw number of reads	Reads after filtering	Remaining % of reads	Library type
Linage I	128,599,724	105,179,065	81%	Paired reads*
Linage IV	124,232,804	110,217,128	88%	
O. lehmanni	102,225,680	91,493,848	89%	
O. sylvatica	48,929,329	45,147,654	92%	Single reads
O. lehmanni	43,633,080	41,086,379	94%	
Linage IV	7,998,636	7,173,285	89%	
O. histrionica	50,204,617	47,376,912	94%	
Reference transcriptome	N/A	33,895,474	100%	Single reads

*. Number of reads correspond to the combined value of both ends libraries.

Table S4A.2. Quality assessment report for each individual and the combined reference transcriptome based on the protein count of the core eukaryotic genes (CEGMA) (Parra *et al.*, 2007; Parra *et al.*, 2009).

Transcriptome	Library	CEGMA mapping	
		Complete	Partial
Linage I	Paired	59%	76%
Linage IV	Paired	60%	85%
O. lehmanni	Paired	63%	85%
O. sylvatica	Single	89%	92%
O. lehmanni	Single	68%	87%
Linage IV	Single	37%	60%
O. histrionica	Single	72%	90%
Reference assembly	Single	83%	93%

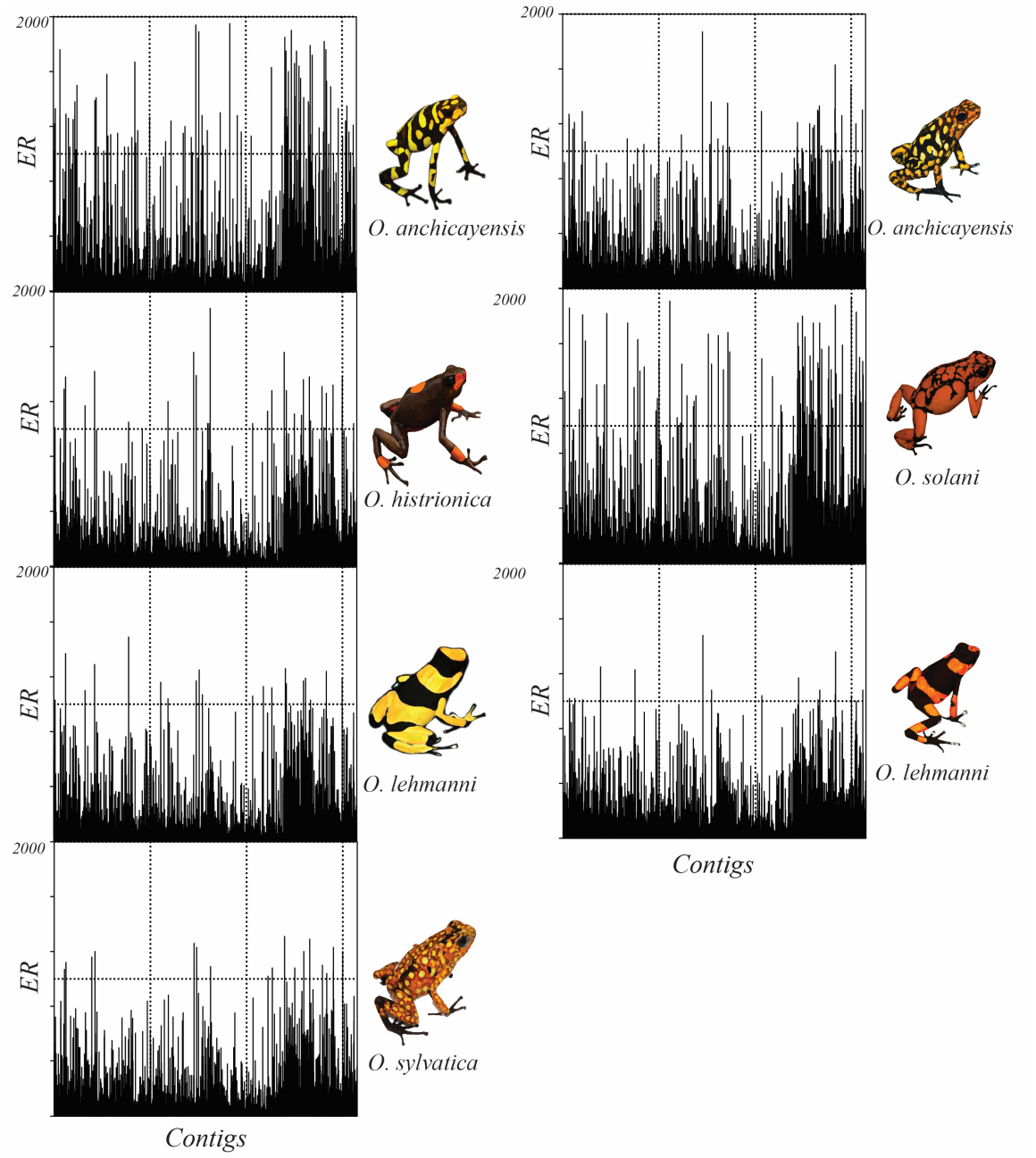


Figure S4B.1. Distribution plots of the effective number of reads (ER) that mapped to the contigs in the reference transcriptome as estimated in eXpress (Roberts & Pachter, 2013).

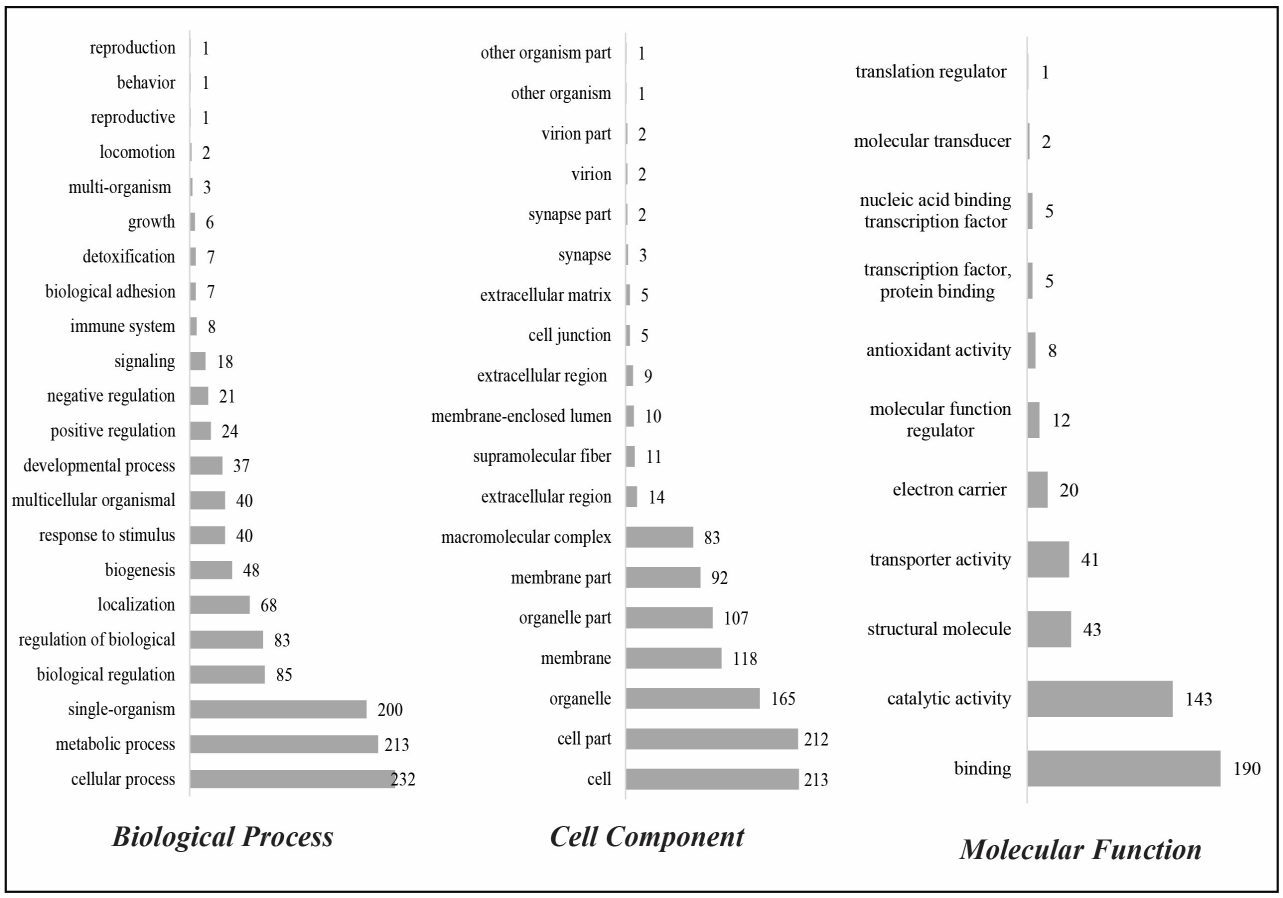


Figure S4B.2. GO mapping distribution (level II) of 1,093 annotated and highly expressed unigenes in the reference *Oophaga* skin transcriptome.

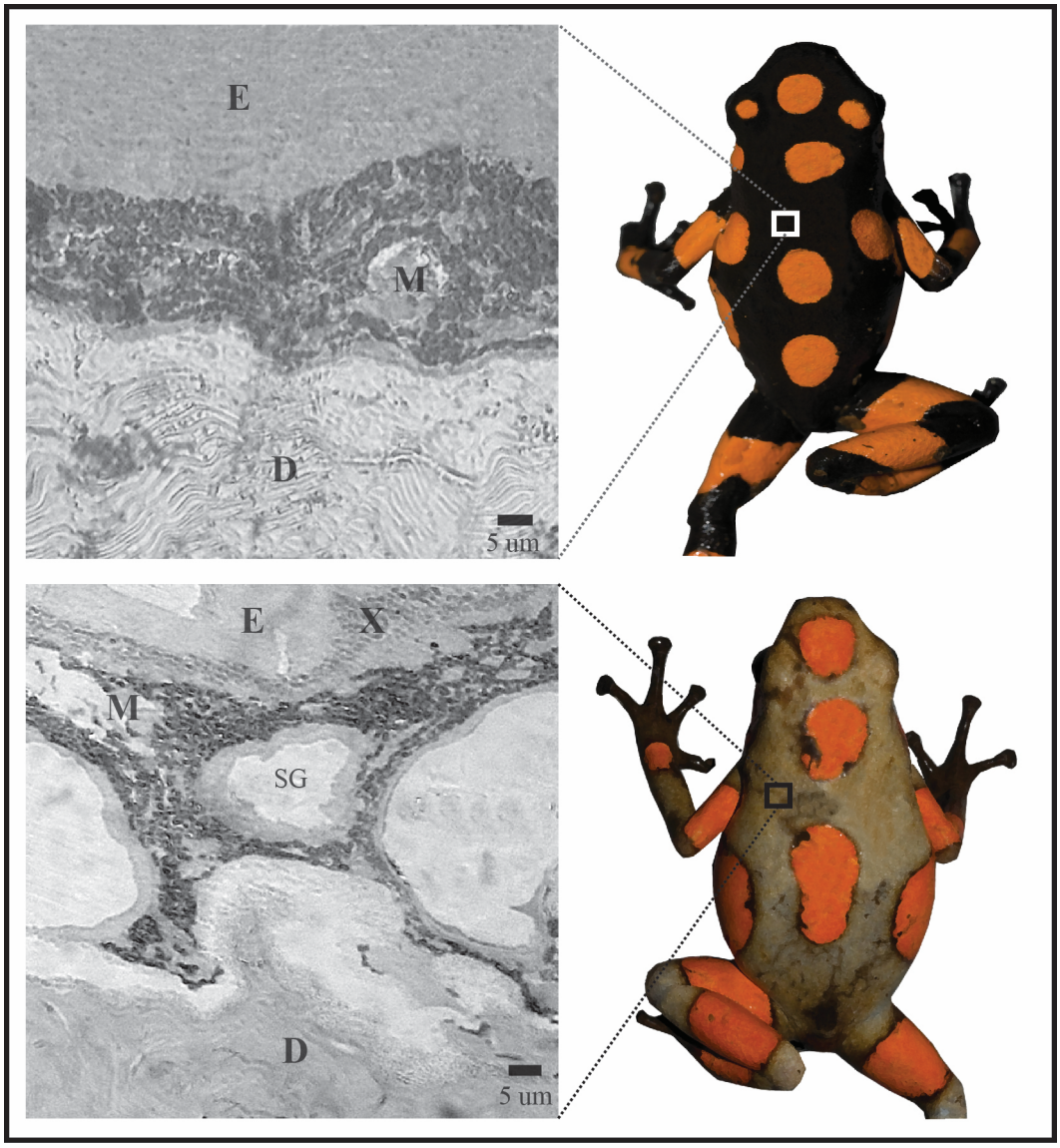


Figure S5A.1. Bright field microscopy (BFM) photographs of contrasting dark-black (upper) and light-brown (lower) background colouration. E=epidermis, M=melanophore layer, D=dermis, X=xanthophores, SG=skin gland.

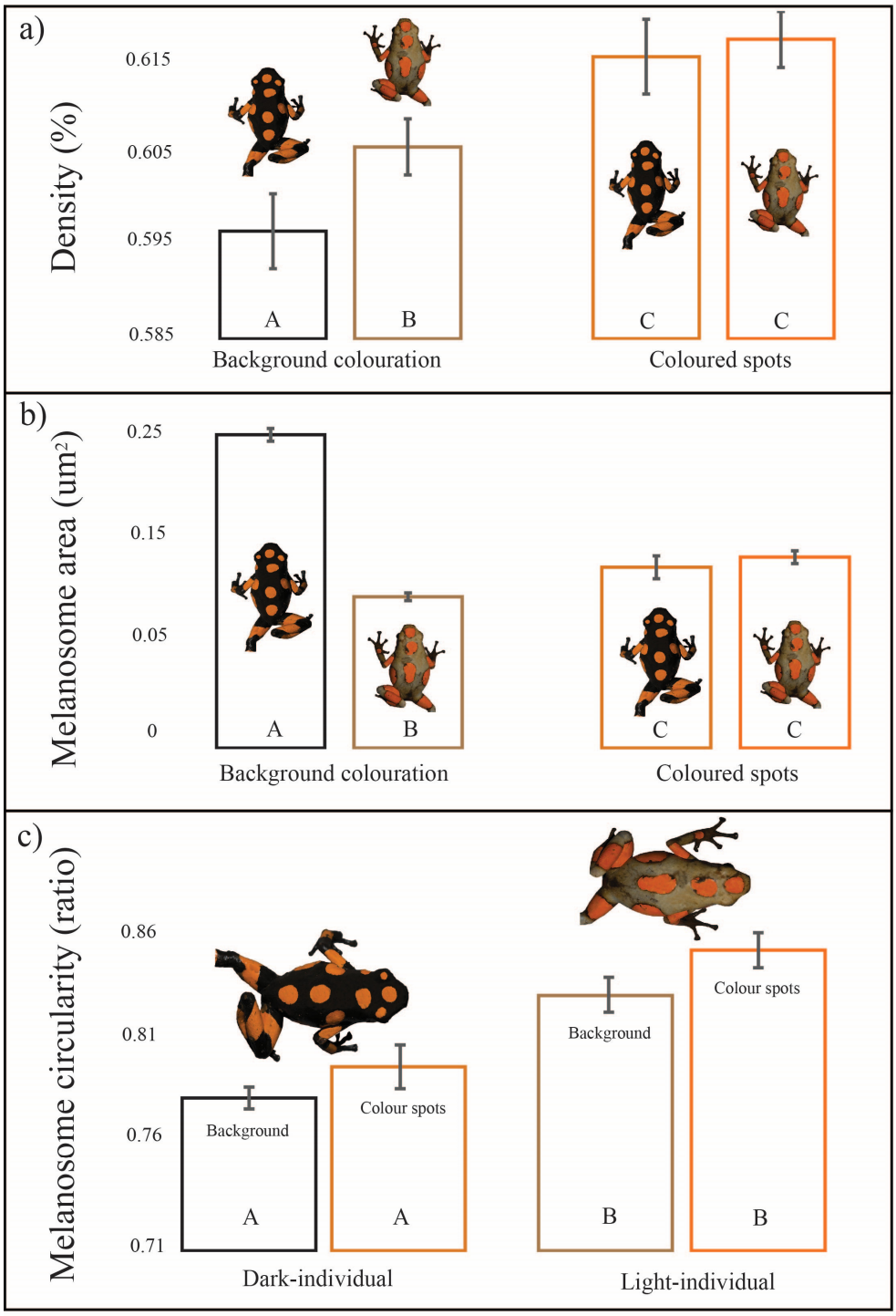


Figure S5A.2. Bar charts representing the mean and standard error value of (a) melanosome density, (b) area, and (c) circularity in four different skin tissues evaluated in this study (see methods).

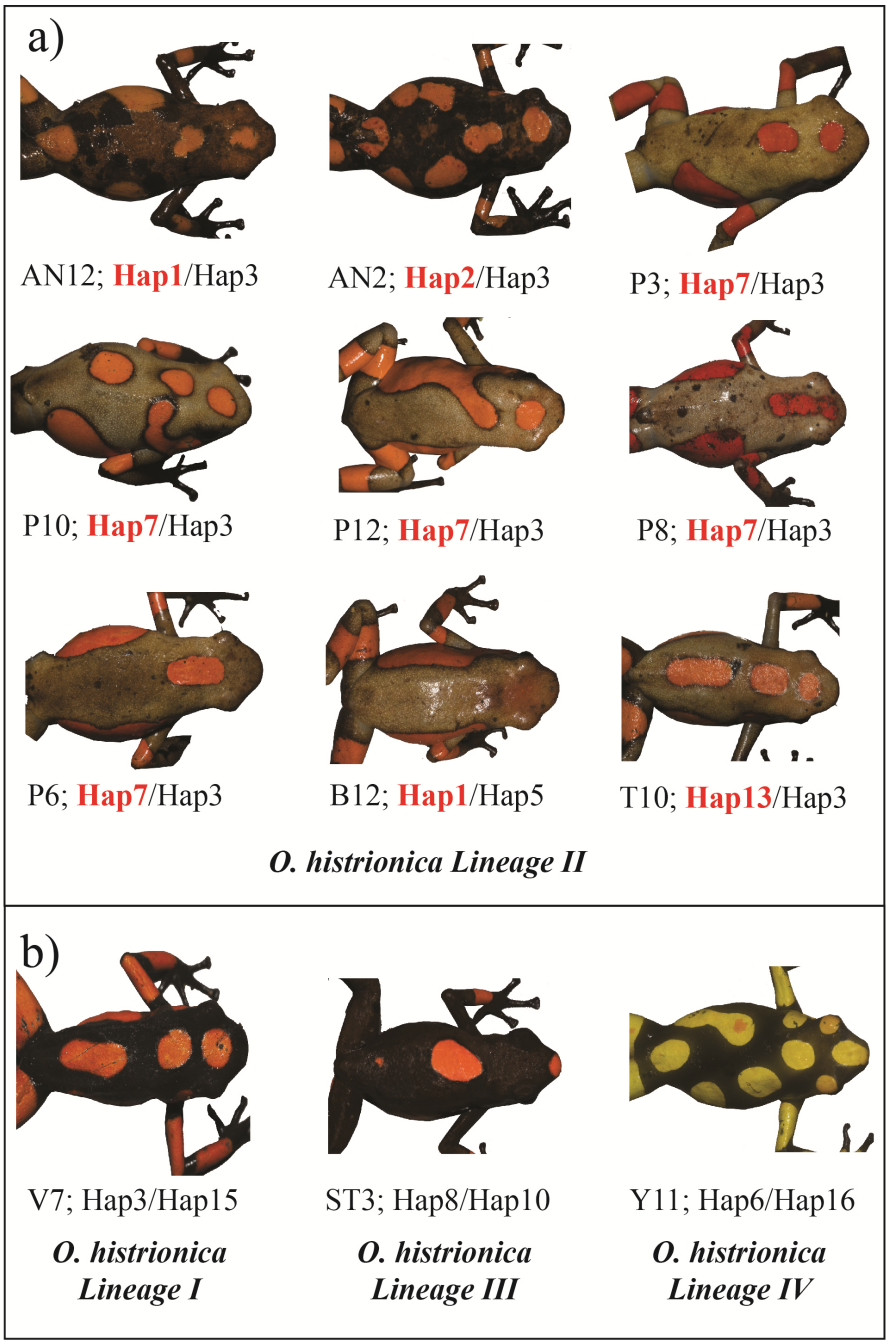


Figure S5A.3. a) Nine heterozygous individuals showing the presence of a "functional" allele (i.e. coding for a 7-transmembrane receptor; in red letters) and highly truncated MC1R (black typed letter) alleles. b) Three heterozygous individuals for alleles coding truncated MC1R receptors. Haplotype number correspond to those in Figure 5.3 and Table 5.4. Lineage number correspond to that described in Fig. 3-1 of Chapter 3. AN=Angostura, P=Pacurita, B=Barranquito, T=Tutunendo, V=Victoria, ST=Santa Cecilia, Y=Anchicaya.

APPENDIX S5B: Supplementary Tables Chapter 5

Table S5B.1. Description of eight morphologic variables used for multivariate estimation of morphological clusters and hierarchical analysis. All variables are measured in RGB units.

Position	Variable	Description of measurement	Units/Type
Dorsal	**Dorsal Background red intensity** (red_b_d)	R value (RGB component) dorsal background colouration	RGB Units
Dorsal	**Dorsal Background green intensity** (green_b_d)	G value (RGB component) dorsal background colouration	RGB Units
Dorsal	**Dorsal Background blue intensity** (blue_b_d)	B value (RGB component) dorsal background colouration	RGB Units
Dorsal	**Dorsal Background Brightness Index** (indexbrigh_ind_b_d)	Dorsal background colouration (brightness index calculated as 0.299R+0.587G+0.114B)	RGB Units
Dorsal	**Dorsal coloured red intensity** (red_d_d)	R value (RGB component) dorsal coloured colouration	RGB Units
Dorsal	**Dorsal coloured green intensity** (green_d_d)	G value (RGB component) dorsal coloured colouration	RGB Units
Dorsal	**Dorsal coloured blye intensity** (blue_d_d)	B value (RGB component) dorsal coloured colouration	RGB Units
Dorsal	**Dorsal coloured Brightness Index** (brigh_ind_d_d)	Dorsal bright colouration (brightness index calculated as 0.299R+0.587G+0.114B))	RGB Units
Dorsal	**Dorsal Brightness difference** (brigh_diff_d)	Dorsal Brightness difference (brigh_ind_d_d) - (brigh_ind_b_d)	RGB Units
Ventral	**Ventral Background red intensity** (red_b_d)	R value (RGB component) ventral background colouration	RGB Units
Ventral	**Ventral Background green intensity** (green_b_d)	G value (RGB component) ventral background colouration	RGB Units
Ventral	**Ventral Background blue intensity** (blue_b_d)	B value (RGB component) ventral background colouration	RGB Units
Ventral	**Ventral Background Brightness Index** (indexbrigh_ind_b_d)	Ventral background colouration (brightness index)	RGB Units
Ventral	**Ventral Brightness difference** (brigh_diff_d)	Ventral Brightness difference (brigh_ind_d_d) - (brigh_ind_b_d)	RGB Units

Table S5B.2: Results from the codon-based Z-tests of neutrality (dN=dS), positive (dN>dS) and purifying selection (dN<dS). Asterisks indicate statistically significant values (P<0.05).

Species	Neutrality		Positive		Purifying	
	Z	**P**	**Z**	**P**	**Z**	**P**
Linage I	-0.98	0.33	-1.00	1.00	1.20	0.12
Linage II	0.68	0.50	0.75	0.23	-0.73	1.00
O. histrionica	1.79	0.08	1.60	0.06	-1.67	1.00
O. lehmanni	0.00	1.00	0.00	1.00	0.00	1.00
Linage IV	2.65	0.01*	2.97	0.00*	-2.54	1.00
Overall	0.40	0.69	0.40	0.34	-0.34	1.00